Title
Essays in Policy Evaluation

Author
Sharma, Himani Vardhan

Publication Date
2024

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Essays in Policy Evaluation

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Economics

by

Himani Vardhan Sharma

Dissertation Committee:
Professor Yingying Dong, Chair
Professor Matthew Freedman
Associate Professor Damon Clark

2024

DEDICATION

To my beloved parents;
My source of strength, love, and joy.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my gratitude towards the individuals who have contributed towards the completion of this dissertation.

I would like to thank my advisor Yingying Dong for her support and guidance during this process. I would further like to thank Matthew Freedman and Damon Clark for their comments and guidance on my work.

I am grateful to Tammy Leonard, co-author of the second chapter of my dissertation, for providing me with the opportunity to work with her data sets and for giving me permission to use our work as a chapter of my dissertation.

This work has also benefited from the Economics department at UCI (during seminars and presentations).

My greatest appreciation goes to my beloved parents. My journey wouldn't have been possible without their constant love and support. I am also thankful to my brother Div for encouraging me throughout this journey.

Finally, I am thankful for the Summer Fellowships from the Economics Department at UCI.

VITA

Himani Vardhan Sharma

EDUCATION

Doctor of Philosophy in Economics **2024**
University of California, Irvine

Master of Arts in Economics **2021**
University of California, Irvine

Master of Science in Economics **2019**
University of North Texas

Post Graduate Diploma in International Business Operations **2017**
Indira Gandhi National Open University

Bachelor of Arts (hons) Economics **2016**
University of Delhi

TEACHING EXPERIENCE

Teaching Assistant **2019-2024**
University of California, Irvine

Teaching Assistant **2017-2019**
University of North Texas

Teaching Fellow **2018-2019**
University of North Texas

FIELDS OF STUDY

Labor Economics, Econometrics, Health Economics

ABSTRACT OF THE DISSERTATION

Essays in Policy Evaluation

by

Himani Vardhan Sharma
Doctor of Philosophy in Economics
University of California, Irvine, 2024
Professor Yingying Dong, Chair

This dissertation contains three chapters. The first chapter attempts to employ a different research design to study the effects of the Deferred Action for Childhood Arrivals (DACA) policy. It employs a cross sectional differences in regression discontinuity design to study the labor market impacts of DACA. The second chapter studies the impact of choice architecture on food selections in the food pantry context. Finally, the third chapter evaluates the efficacy of a conditional cash transfer in improving the status of girls in India.

In the first chapter, I study the impacts of Deferred Action for Childhood Arrivals (DACA) on individuals' labor market outcomes using a cross sectional differences-in-regression-discontinuity (DRD) design. The DRD design leverages the DACA multi-dimensional eligibility criteria. In particular, in order to be eligible for DACA, one has to be born after a certain cutoff date, among others. The DRD design then leverages this age discontinuity to create two RD designs, one among those potentially DACA eligible individuals, while the other counterfactual RD design among those DACA ineligible individuals (by other DACA eligibility criteria). The DRD design is valid under weaker conditions than the standard RD framework. Using data from the American

Community Survey (ACS), I find that the DACA eligible population earned a higher income. I also find suggestive evidence of an increase in labor force participation.

The second chapter highlights the impact of choice architecture, and in particular relative trade-offs, on food selections in the food pantry context. Client choice food pantries allow individuals, many of whom are food insecure, to select a preferred bundle of food. To date, interventions to improve the nutrition of food choices in pantries have not included price incentive programs like those employed in the retail food sector because pantries do not charge for food. However, economic incentives may still play a role in food pantry choices through choice architecture. We examined a natural experiment involving two client-choice regimes that effectively altered the opportunity cost of food selections. Longitudinal individual fixed effects models provide evidence that pantry clients responded to changed opportunity costs by selecting more foods that became relatively less expensive and fewer foods that became relatively more costly.

The third chapter sheds light on the effectiveness of long-term cash transfers in improving the status of women in India. In response to the prevalence of female feticide, Bihar launched a policy called Mukhyamantri Kanya Suraksha Yojana (MKSY) in 2008. This policy aimed to improve the social status of women and to improve the sex ratio in the state. The policy provided long term cash transfers to two daughters of a family if certain eligibility conditions were satisfied. I analyze if the policy led to an improvement in the survival rate of the girl child. Furthermore, I also study the effects of the policy on the schooling outcomes of the girl child. My study finds negligible effects of the policy.

Chapter 1

Does DACA Affect Labor Market Outcomes? Evidence from a Cross Sectional Differences in Regression Discontinuity Design

1.1 Introduction

The United States is home to a large immigrant population and consisted of 44.8 million foreign born population in 2018.[1] Many discussions concern those who are present in the country without paperwork and hence, do not have legal status. This unauthorized population lives under a constant threat of deportation. Furthermore, due to their undocumented status, they have less access to employment opportunities as compared to the authorized sections of the population.

This paper studies the policy of Deferred Action for Childhood Arrivals (DACA) introduced in 2012. The policy was introduced to allow certain unauthorized people to stay in the US and provide work authorization to them. In a report by Brookings in 2013, through 2013, more than half a million people applied for DACA. As of March 31, 2021, there are approximately 616,030 active DACA recipients and about 80.6% of them are Mexicans.[2]

The population eligible for the policy has witnessed improved health, labor, and schooling outcomes. For example, studies find an increase in hours of work and work probability for the DACA eligible population (Pope (2016)) and reduced poverty among households (Catalina

[1] Pew Research Center
[2] https://www.uscis.gov/sites/default/files/document/reports/Active%20DACA%20Recipients%20%E2%80%93March%2031%2C%202021.pdf

Amuedo-Dorantes (2016)). DACA eligible population is also less likely to delay care because of financial constraints (Guintella, Lonsky (2020)) and has improved health outcomes (Patler, Caitlin et al. (2019)), improved health insurance rates (Jung (2020)), and positive mental health impacts (Venkataramani et al. (2017)).

There exists a small literature evaluating the labor market impacts of DACA, but these papers vary in data/research designs used and reach different conclusions regarding the impacts of DACA. Therefore, it is still an unsettled question as to whether DACA improves labor market outcomes for the eligible population. My paper utilizes a novel research design that leverages the multiple eligibility criteria of the DACA policy and uses a differences-in-regression-discontinuity (DRD) methodology in a cross-sectional framework, which is valid under weaker conditions than the standard RD design. The standard RD design has been considered as being highly credible and warrants strong internal validity. My research design improves upon the standard RD design. Findings from my paper thereby add additional compelling evidence on the potential impacts of DACA on education and labor market outcomes.

To be eligible for DACA, one has to be born after a certain cutoff date, among others. The DRD design leverages this age discontinuity to create two RD designs, one among those potentially DACA eligible individuals, while the other counterfactual RD design among those DACA ineligible individuals (by other DACA eligibility criteria). The DRD design relaxes the assumption of continuity of pre-determined variables at the cutoff (a standard assumption required for the validity of the RD design).

Using data from the American Community Survey (ACS), I find that the DACA eligible population earned a higher income. The estimates range from $4,168 to $6,016 per year. I find that this increase in income is due to an increase in wages. I also find suggestive evidence of an increase in weekly hours of work by around two hours.

Using the policy's multiple eligibility criteria, I test the robustness of my results to alternative counterfactuals. I exploit the multiple eligibility conditions of the DACA policy to create different control groups. Therefore, these alternative counterfactuals are composed of different groups of the ineligible population. I further test the robustness of my results to a variety of functional forms (linear and quadratic specifications). My results add to the current mixed evidence of DACA on labor market outcomes (Pope, 2016, Ameudo-Dorantes Antman, 2017, Hamilton et al, 2021).

The rest of the paper progresses as follows. Section 2 describes the policy. Section 3 describes the literature. Section 4 describes the data and the sample creation procedure. Section 5 discusses the empirical strategy and the econometric specification. Section 6 discusses the results. Section 7 concludes the study.

1.2 Deferred Action for Childhood Arrivals (DACA)

DACA was introduced by former President Barack Obama in 2012. It provided temporary relief from deportation to a certain unauthorized population who entered the US at a young age. In addition, it provided work authorization.

In order to meet the eligibility requirements, the applicant should have been born after June 15, 1981; should have entered the United States before they reached their 16th birthday; should have been continuously residing in the United States since June 15, 2007 up to the present time; was physically present in the United States on June 15, 2012 and at the time of making request for DACA with the USCIS; had no lawful status on June 15, 2012; is currently in school, or has completed high school, or has gotten a GED certificate, are honorably discharged veteran of the

Coast Guard or Armed Forces of the United States; and is not guilty of a criminal conduct or a threat to the national security. In addition, the applicant must be at least 15 years old.[3]

Despite its potential positive effects, the policy started to face challenges within a few years of its introduction. Then President Donald Trump rescinded the policy in September 2017 and only allowed application renewals. President Joe Biden reinstated the policy to its previous eligibility (of 2012) in 2021, but a Texas court invalidated it again, allowing only application renewals.

1.3 Literature Review

In this section, I discuss some existing studies on DACA and summarize my contributions to the literature. Some research on DACA has focused on labor market outcomes of DACA eligible individuals. This research offers mixed evidence regarding the impacts of DACA on different labor market outcomes. For example, Pope (2016) found an increase in income for the eligible population. His study found an increase in income at the lower end of the income distribution. Hamilton et al (2021) found no impact of DACA on the labor force participation rates. Ameudo-Dorantes Antman (2017) found that DACA increases the likelihood of employment. These studies generally employ a difference in differences (DID) approach to study the effects of the policy. Contrary to the common DID approach in the DACA literature, I conduct a study on the labor market impacts of DACA using a differences-in-regression-discontinuity (DRD) design in a cross-sectional framework and offer additional contribution to the DACA literature through a different research design.

I use data from the American Community Survey (ACS) and exploit the age eligibility condition of being under the age of 31 as of June 15, 2012, that is, being born after June 15, 1981, as a source

[3] Conditions directly taken from the USCIS guidelines.

of discontinuity among the potentially DACA eligible population. Since the exact date of birth is not available, the RD design in this paper uses quarter of birth as a running variable.

This study is closest to Pope (2016) in terms of the data set used and the outcomes studied. Pope (2016) used data from the American Community Survey (ACS) for the years 2005-2014. He used a DID analysis and exploited variation in age eligibility (in years), the age of entry of the individual in the US, and subsequently the entire variation in DACA eligibility to estimate the impact of DACA.

This study is also closely related to the study of Jung (2020) in terms of the methodology used. Jung (2020) used data from the ACS and showed that DACA led to an increase in health insurance among DACA eligible compared to those DACA ineligible. The study, using a DRD, explored the difference between pre and post-policy implementation and used variation from the eligibility condition of the age at which an individual entered the United States. The model made use of a more credible research design than the DID specification.

By applying cross sectional differences in regression discontinuity design and using quarter and year of birth as the running variable, I provide additional evidence on the efficacy of DACA on the labor market outcomes of DACA eligibles. My research design improves the standard RD design. As I show in the empirical specification section (section 1.5), contrary to the requirement of the continuity of the pre-determined covariates in a standard RD design, my model allows for a jump in pre-determined variables at the cutoff as long as the jump is similar between the treatment group and the counterfactual groups. The DRD design therefore requires weaker assumptions than the standard RD design and allows for possible discontinuity in the pre-determined covariates (which is important in this context, as I discuss in greater detail in section in section 1.5).

1.4 Data

In this section, I summarize the data used in the study and the sample creation process.

I use individual level data from the American Community Survey (ACS) available at Integrated Public Use Microdata Series (IPUMS)-USA. It is one of the largest surveys conducted in the United States and collects information regarding demographics, occupation, etc. ACS covers a representative sample and has a high rate of response, making it a good source of obtaining estimates, which are nationally representative.

The data consideration for this study ranges from 2014 to 2017 and I study the labor market impact of DACA on the total income, probability of work, weekly hours of work, and probability of current school attendance.[4] I use the year 2014 as the first year in the post-DACA period since all the three outcome variables pertaining to an individual's labor decisions refer back to the past 12 months, that is, if the applicant was surveyed in February of the present year, the questions refer to the period ranging from the month of February in the present year to the month of March in the past year. Since the USCIS started accepting applications in August of 2012, considering the acceptance of applications and the processing time, taking the first period as 2013 would lead to a reference towards the months in 2012 when the impact of DACA was likely small. The period of 2017 is taken as the final post DACA period since then President Donald Trump rescinded DACA in September 2017, which did not allow for any new acceptances of the DACA applications. Hence, the eligibility criteria for DACA changed from September 2017.

To create the sample for the actual RD (in contrast to the sample for the counterfactual RD), I use the several conditions required to be eligible for DACA. I use the questions in the ACS survey to

[4] School attendance variable measures the likelihood that the person attended any school leading to a high school or a college degree in the past three months.

create the treatment group, which satisfies the above DACA eligibility conditions, and the process is similar to that of Pope (2016).[5] The treatment group consists of population that satisfies all the DACA eligibility conditions (that can be identified from the data) except for the age requirement of being born before or after June 15, 1981, leading to a discontinuity in potential eligibility arising from the age requirement.

Suppose one observes all the variables that can be used to determine DACA eligibility. Then a standard sharp RD design can be used to estimate an intent to treat effect (ITT), i.e., the impact of DACA eligibility on the outcomes of interest. Under this scenario, the DACA eligible population and the DACA ineligible population can be perfectly identified which would allow for the computation of the exact intent to treat effects.

Due to data limitations, I cannot observe all the conditions required for DACA eligibility. For instance, ACS does not provide any information regarding the criminal status of an individual or the true legal status of an individual. Therefore, I create the sample based on the information available, which includes citizenship, year of immigration, educational qualification and the year and quarter of birth of individuals.

I create the treatment group using four steps. Firstly, I make use of the citizenship information in the ACS and include only the individuals who I believe could be DACA eligible and are the largest applicants of DACA. As my second step, I use the year of immigration question in the ACS to keep only the individuals I selected in the first step who satisfy the year of immigration requirement of residing for at least 5 years before the policy announcement. Following this, I exclude individuals who came to the United States after the age of 16 years. Finally, of the sample I created

[5] Code from the study is publicly available and has been used for data cleaning in this study.

in the third step, I keep only the individuals who have at least a high school degree. These four steps lead to a selection of a sample where every individual is potentially eligible for DACA. This is followed by exploiting the discontinuity based on the date of birth. I elaborate on this process below.

One of the conditions require that the applicant had no lawful status on June 15, 2012. The ACS does not ask whether an individual has legal status or not. Hence, I use the question on citizenship to proxy for legal status and include only non-citizens in the population. Furthermore, following the common approach in DACA literature, I include only Mexican non-citizens since Mexicans are the largest applicants of DACA. Hence, both sides of the cutoff are potentially eligible based on citizenship information.

The next condition requires me to find out if the individual has continuously resided in the United States for at least five years on June 15,2012. The ACS asks the year of immigration of all the individuals. From the condition, anyone who immigrated after June 15, 2007, will be ineligible for DACA. Since the questionnaire does not ask the exact date of immigration, I exclude all individuals who immigrated in 2007 or after.

DACA also takes into account whether an individual entered United Sates below the age of 16 years. ACS asks applicants the number of years they have been residing in the United States. This question along with the age of the applicant is used to find out the date of entry in the United States. I keep individuals who entered the US below the age of 16.

Next, the eligibility criteria require me to find out if the applicant satisfies particular educational requirements, which can be found out by the education information offered by the survey. In order to satisfy the education requirement, I include Mexican non-citizens with at least a high school degree.

Finally, the eligibility condition requires me to find out if the applicant was under the age of 31 as of June 15, 2012, that is, if the applicant was born after June 15, 1981. While ACS does not provide the exact date of birth for an individual, it provides information regarding the quarter and year in which the person is born. I exclude the people born in the year 1981 and in quarter 2 (since I cannot observe their exact eligibility). Following this, people born in 1981 (quarter 3) or afterwards will be potentially eligible for DACA while people born in 1981 (quarter 1) or before will become ineligible for DACA due to failure of meeting age requirements. I consider the bandwidths of 16, 20, and 24 quarters around this cutoff.

Table 1.1 shows the descriptive statistics of the population for the years 2014-2017 for the control group and the treatment group above and below the cutoff of age eligibility for the bandwidth of 20 quarters.[6] [7] The control group consists of Mexican population, which is ineligible for DACA based on other eligibility criteria other than the criterion of age eligibility. The first column shows the summary statistics for the treatment group (potentially DACA eligible population). The second column shows the summary statistics for the control group (DACA ineligible population). The third column shows the difference between the means of these two groups. The fourth column shows the *t*-statistic arising from the test for the difference between the two means. The clearest differences between the potentially DACA eligible group and the DACA ineligible group is that the potentially DACA-eligible group tends to have entered the United States at a younger age.

[6] Binary variables are coded in percentage terms; Observations "above cutoff" refer to the sample that is born after June 15, 1981 (hence, DACA eligible based on the age discontinuity condition) and observations "below cutoff" refer to the sample that is born before the cutoff of June 15, 1981 (hence, DACA ineligible based on the age discontinuity condition). The treatment group is defined above. The control group consists of Mexican population, which is ineligible for DACA based on other eligibility criteria other than the criterion of age eligibility.
[7] Appendix (A) Tables A1 and A2 show similar tables for bandwidths 16 and 24.

Table 1.1: Descriptive Statistics (Bandwidth:20 Quarters)

Variable	Mean		Difference	t-statistic
	Treatment Group	Control Group		
Treatment Group				
Total Income (in $1000)	25,925.02	27,950.91	-2,025.89	-4.89
Work	0.82	0.79	0.03	4.68
Hours of Work	32.44	31.78	0.66	2.62
School Attendance	0.06	0.06	0.00	1.32
Male	0.54	0.52	0.02	3.36
Age of Entry in the US	8.77	19.01	-10.24	-133.35
Cognitive Difficulty	0.02	0.03	-0.01	-6.89
Ambulatory Difficulty	0.01	0.02	-0.01	-5.05
Mobility Difficulty	0.01	0.02	-0.01	-5.50
Self-Care Difficulty	0.01	0.01	-0.00	-2.46
Hearing/Seeing Difficulty	0.02	0.02	-0.00	-2.25
Observations	7,469	155,477		

1.5 Empirical Strategy

In this section, I elaborate on the empirical specification used in this study. The aim of my analysis is to study the impact of DACA on the labor market outcomes namely, total income, probability of working and the weekly hours of work. In addition, I also study the impact of DACA on the likelihood of being in post-secondary school to see if the policy impacted schooling decisions. I conduct the analysis at bandwidths of 16, 20, and 24 quarters of data at either side of the cutoff (the chosen bandwidths allow me to study the population between the ages of 27 and 42). I first use a standard RD design to conduct my baseline analysis and then extend my results using the DRD design. I take the age of individual on June 15, 2012, as the cutoff deciding the discontinuity in potential eligibility for DACA.

For my baseline analysis using the treatment group, individuals below the cutoff consists of the individuals in the treatment group who are born before June 15, 1981, and the ones above the cutoff denote the ones who are born after June 15, 1981. The treatment is being eligible for DACA policy. Since I cannot observe the exact date of birth of an individual, the running variable is the normalized year and quarter of birth of an individual.

This gives rise to the reduced form model as described in equation (1). I choose a linear specification and allow the slope of the function to vary at either side of the cutoff.[8]

$$Y_i = \delta_0 + \delta_1 R_i + \delta_2 D_i + \delta_3 R_i D_i + \epsilon_i \quad (1)$$

where R_i refers to the running variable, that is, the normalized year and quarter of birth of an individual; D_i indicate whether one is eligible for DACA based on the age criterion for DACA. The analysis is performed using robust standard errors and is weighted using sampling weights.

The parameter of interest is δ_2. Since I do not observe all the DACA eligibility conditions from the data, δ_2 does not measure precisely the intent to treat (ITT) effects of DACA. In particular, because I approximate some of the eligibility criteria in my sample, the sample inevitably includes ineligibles individuals on both sides of the RD cutoff, which attenuates the ITT.[9]

I then extend the baseline analysis using the DRD approach where I now include a control group (consisting of DACA ineligibles based on criteria other than the age criterion) in the analysis.[10] Specifically, I now use the entire Mexican post DACA sample and create a dummy variable, S, a treatment group indicator. S =1 if an individual belongs to the treatment group defined previously and S =0 if an individual belongs to the control group, i.e., the group that is not eligible for DACA based on other DACA criteria other than the age criterion.

[8] I show the robustness of results using alternative specifications as the study progresses.

[9] For instance, consider documented/undocumented status. Since I cannot observe whether an individual is unauthorized/undocumented or not, I restrict the treatment group to Mexican non-citizens. Hoefer et. al (2012) estimate that there were around 6,800,000 unauthorized immigrants between the ages of 25 to 44 years in the US in January 2011.This age group includes the subset of individuals (in terms of age) considered in my study. Furthermore, for this age group, I compute the non-citizen population from the 2010 ACS. The statistic was computed to be around 11,138,841. Hence, the unauthorized population makes up around 61% of the non-citizen population. Since the age group considered in my analysis is a subset of this age group, the sample in this study will be at least 61 % unauthorized. Furthermore, due to the inclusion of non-citizens who are only Mexicans, the unauthorized population will make up more than 61% of the non-citizen population since Mexican non-citizens form a major portion of this group.

[10] An alternative to doing the DRD specification is to include quarter fixed effects in the baseline RD analysis. Although, this is a possibility, the DRD methodology is preferred due to the higher extrapolation required since the running variable is discrete. Furthermore, the DRD methodology will lead to a differencing out of the quarter effects (at the cutoff) assuming that the confounding effects are similar across groups. I show the RD results with quarter fixed effects in the Appendix (A) (Tables A8). The results do not show much deviance from the DRD results.

Equation (2) gives the reduced form equation where the rest of the variables carry the same meaning as in equation (2) and the coefficient of interest is the interaction term between S and D.

$$Y_{isc} = \beta_0 + \beta_1 D_i + \beta_2 S_i + \beta_3 S_i D_i + \beta_4 R_i + \beta_5 R_i D_i + \beta_6 R_i S_i + \beta_7 R_i S_i D_i + \epsilon_{isc} \qquad (2)$$

where the parameter of interest is $\beta 3$, which represents the difference between the ITT for the treatment group and the counteract ITT for the control group. I use two sources of variation to compute the effects, that is, the variation from the treatment group assignment across S and the variation from across the quarter-year cohort, c.

In order to show what the model in equation (2) is estimating, I compute Equation (2) separately for two groups, one pertaining to the control group and the other to the treatment group.

When $S_i = 0$ (that is, for observations in the control group), equation (2) becomes:

$$Y_{isc} = \beta_0 + \beta_1 D_i + \beta_4 R_i + \beta_5 R_i D_i + \epsilon_{isc} \qquad (3)$$

When $S_i = 1$, (that is, for observations in the treatment group), equation (2) becomes:

$$Y_{isc} = (\beta_0 + \beta_2) + (\beta_1 + \beta_3) D_i + (\beta_4 + \beta_6) R_i + (\beta_5 + \beta_7)\, R_i D_i + \epsilon_{isc} \qquad (4)$$

In equations (3) and (4), the coefficient on Di captures the ITT for the control group and the treatment group, respectively. Therefore, contrary to the requirement of the continuity of the pre-determined covariates (at the cutoff) in a standard RD design, my model allows for a jump in pre-determined variables at the cutoff as long as the jump is similar between the treatment group and the counterfactual groups. β_3, the parameter of interest, is the difference between the two ITTs, which differences out any mean differences in the outcomes that arise purely due to the effects of quarter of birth.

1.6 Results

First, I begin by employing the RD method. Any potential discontinuity derived from the RD method might just be a result of differences arising from the quarter of birth. Hence, I subsequently adopt the DRD methodology to difference out the differences arising due to the quarter of birth differences. I further provide robustness checks to show the robustness of my results.

1.6.1 Validity Check for the Baseline RD

In this section, I check the validity of the baseline RD design. This validity is checked through the continuity of pre-determined covariates at the cutoff. The continuity of covariates through the cutoff ensures that the condition of date of birth is the only thing affecting the outcome variables that is changing at the cutoff. The continuity of pre- determined covariates is tested through regression analysis. This is done by substituting the covariates as outcome variables in the main regression models.

In order to test for the smoothness of covariates across the cutoff, I test the gender of the individual, the age at which an individual entered the US, and several disability variables (that can arise due to quarter of birth differences -- whether the individual suffered from cognitive difficulty, ambulatory difficulty, self-care difficulty, independent living difficulty, and/or hearing/vision difficulty).

Table 1.2 shows the results for the two pre-determined covariates, gender and age of entry in the US, at different bandwidths. I see a weak discontinuity in gender arising at the bandwidth of 20. Furthermore, Appendix (A) figures A1-A6[11] show the continuity of these two covariates at the

[11] I use the code for graph plotting from Clark and Royer (2013).

cutoff across all bandwidths. For the covariate, age of entry, the discontinuity is insignificant across all bandwidths.

Table 1.2: Estimates for Continuity of Covariates (for the Baseline RD)

Outcome Variable:	Gender			Age of Entry in the US		
Estimate	0.034 (0.034)	0.052* (0.031)	0.029 (0.028)	-0.163 (0.335)	-0.175 (0.296)	0.137 (0.270)
Bandwidth	16	20	24	16	20	24

Note: All estimates are weighted using sampling weights; Robust standard errors in parentheses. The analysis is performed using equation (1) . * p < 0.1, ** p < 0.05, *** p < 0.01

Appendix (A) table A3 shows the results for the disability characteristics at different bandwidths. Furthermore, Appendix (A) figures A7-A21 show the continuity of the covariates at the cutoff across all bandwidths. There seems to exist no evidence of a difference between people born in the first quarter and those born in the third quarter. Even though I do not see any significant discontinuities in the covariates available, there might still be discontinuities in the covariates that are not available in the data set. These might confound the results and hence, the DRD will potentially help remove these confounding factors.

1.6.2 Baseline RD Analysis

I first analyze the outcome variables graphically to expose any potential discontinuities arising at the cutoff of being born before or after June 1981. Figures 1.1-1.4 show the discontinuity arising in total income, probability of work, weekly hours of work, and probability of current school attendance at a bandwidth of 20 quarters, respectively.[12] The estimated discontinuities are positive and imply towards an increase in magnitude after the cutoff, that is, individuals who are DACA eligible have a higher income, a greater probability of work and school attendance, and work more hours in the post DACA period. For income, probability of work, and hours of work, the

[12] The graphs for variables at alternative bandwidths of 16 and 24 are shown in the appendix (A) figures A22-A29

discontinuity is significant across all bandwidths. Furthermore, the discontinuity for probability of school attendance is insignificant across all bandwidths.

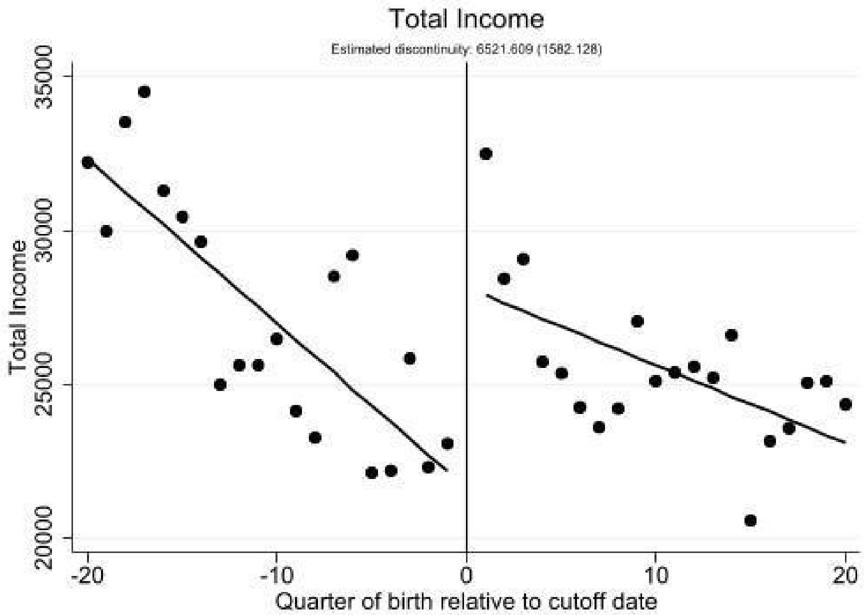


Figure 1.1: Total Income of the treatment group in the post DACA period (2014-2017). The estimated discontinuity is computed using linear regression; robust standard errors in parenthesis; the fitted values from the regression are plotted via lines.

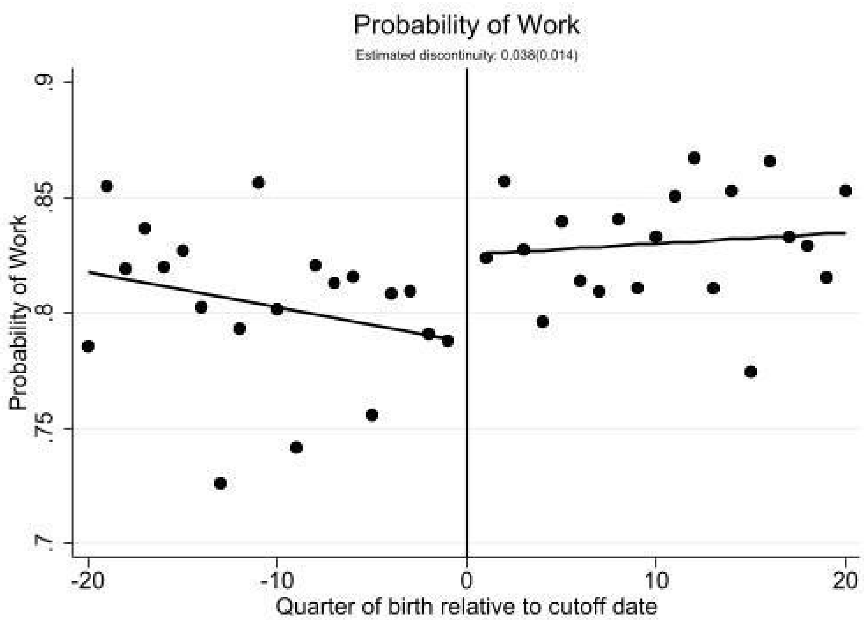


Figure 1.2: Probability of Work of the treatment group in the post DACA period (2014-2017). The estimated discontinuity is computed using linear regression; robust standard errors in parenthesis; the fitted values from the regression are plotted via lines.

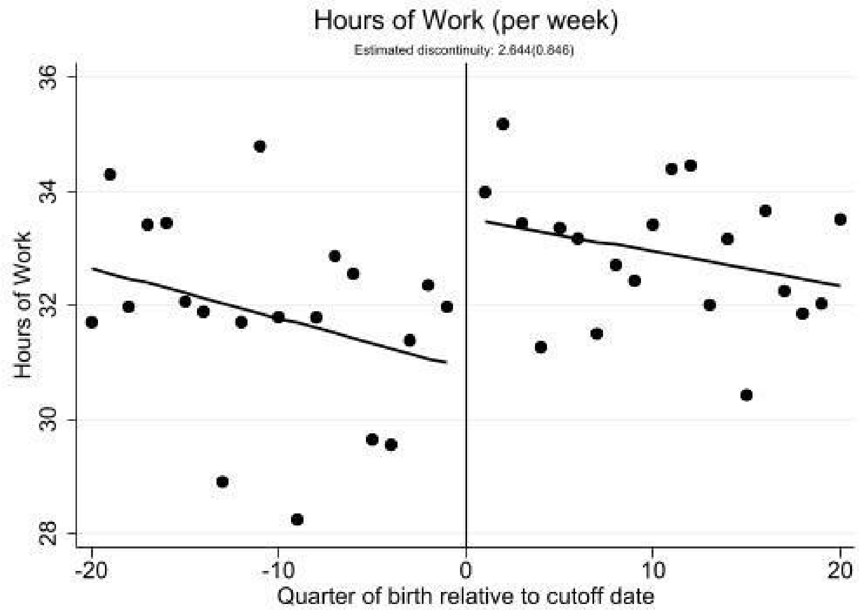


Figure 1.3: Hours of Work of the treatment group in the post DACA period (2014-2017). The estimated discontinuity is estimated using linear regression; robust standard errors in parenthesis; the fitted values from this regression are plotted via lines.

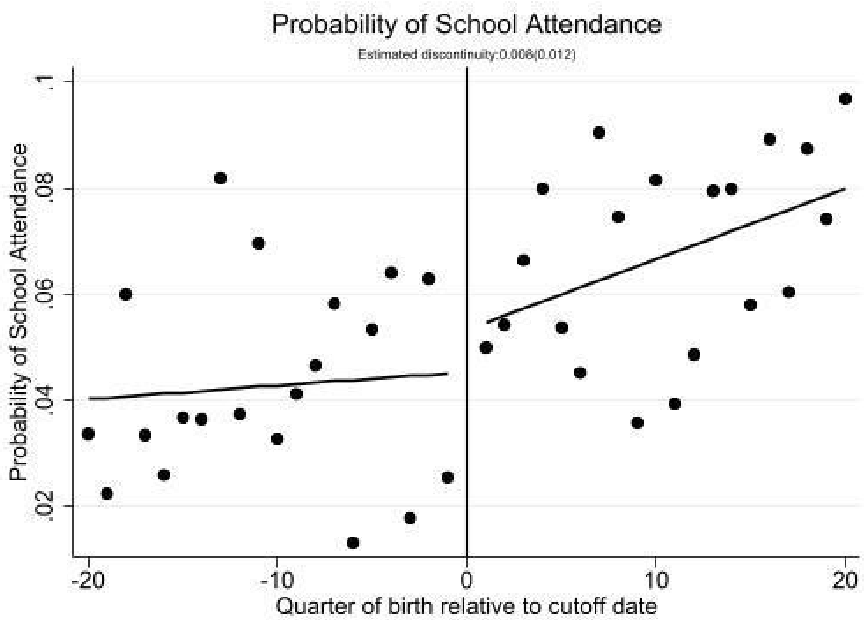


Figure 1.4: Probability of School Attendance of the treatment group in the post DACA period (2014-2017). The estimated discontinuity is computed using linear regression; robust standard errors in parenthesis; the fitted values from the regression are plotted via lines.

Table 1.3 shows the baseline RD analysis results for the outcome variables. The estimates show an increase in total income for people who satisfy the eligibility date, and the results are robust across different bandwidths (in terms of sign and significance). The positive impact ranges from

$4,439 to $6,736, implying an increase in financial stability of people in the treatment group born after the cutoff. These estimates point towards an increase of as much as 25% of the mean income of the sample.

Table 1.3: Baseline RD Results

Outcome Variable:	Total Income			Work or not			Hours of Work			School Attendance		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Estimate	6736.099*** (2029.848)	6259.914*** (1786.314)	4438.892*** (1536.735)	0.033 (0.027)	0.037 (0.023)	0.030 (0.021)	2.478** (1.251)	2.622** (1.093)	1.779* (1.002)	0.012 (0.013)	0.006 (0.012)	0.007 (0.011)
Bandwidth	16	20	24	16	20	24	16	20	24	16	20	24
Mean	25,650.63	25,925.02	25,795.67	0.817	0.820	0.819	32.419	32.442	32.418	0.057	0.059	0.061
Observations	5,731	7,469	9,224	5,731	7,469	9,224	5,731	7,469	9,224	5,731	7,469	9,224

Note: All estimates are weighted using sampling weights; Robust standard errors in parentheses The analysis is performed using (2) . * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Furthermore, I see no evidence of DACA affecting the work probability of people who are born after the cutoff. However, I witness an increase in the weekly hours of work. This indicates an increase in labor force participation in the form of increased hours of work. Additionally, I do not witness any impact on the likelihood of an individual currently attending school.

1.6.3 DRD Analysis

I now extend the baseline analysis using the DRD framework. The DRD analysis is valid under weaker conditions. The baseline RD analysis studies the differences in labor market outcomes of the treatment group by exploiting the discontinuity in potential eligibility arising from the date of birth. Since the exact date of birth is not available, the quarter of birth has been used. However, there could be a possibility that the differences in outcomes are arising due to the quarter of birth of an individual, making the estimates biased. DRD analysis can allow for such confounding factors that invalidate the comparability of observations right above and right below the RD cutoff.

Similar to the process done for the treatment group in the baseline RD analysis, I first analyze the outcome variables in the control group (graphically) to expose any potential discontinuities arising at the cutoff of being born before or after June 1981. To reiterate, the treatment group consists of the Mexican population that satisfies all the DACA eligibility conditions (that can be identified from the data) except for the age requirement of being born before or after June 15, 1981, which leads to a discontinuity in potential eligibility arising from the age requirement. The control group consists of the Mexican ineligibles who are not eligible for DACA based on other DACA eligibility criteria other than the age criterion.

For the control group, Figures 1.5-1.8 show the discontinuity arising in total income, probability of work, weekly hours of work, and probability of current school attendance at a bandwidth of 20 quarters, respectively.[13] The estimated discontinuities are insignificant for probability of school attendance across all bandwidths. I see a significant negligible discontinuity for probability of work at all bandwidths, and a significant positive discontinuity for hours of work at bandwidths of 20 and 24. I also witness a discontinuity amounting to $710 post cutoff for income at a bandwidth of 16.

[13] The graphs for variables at alternative bandwidths of 16 and 24 are shown in the appendix (A) figures A30-A37.

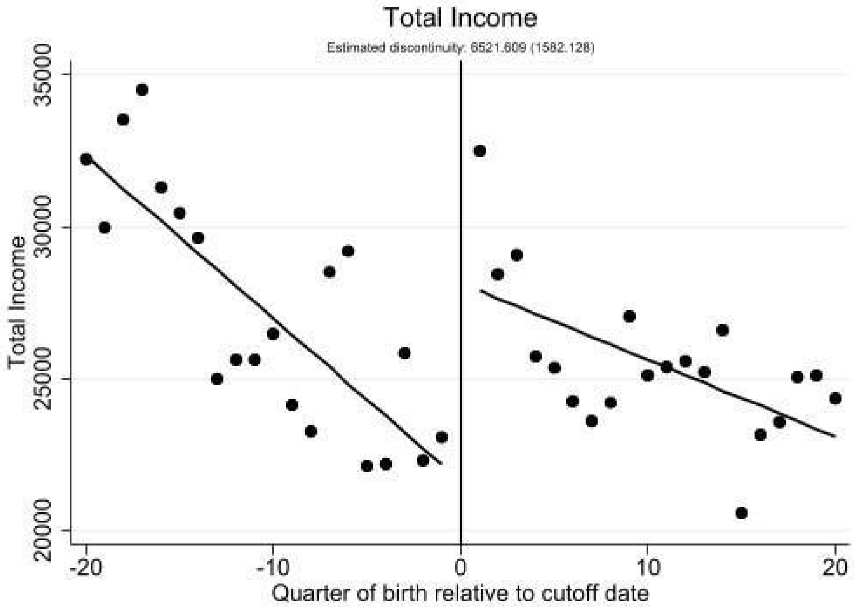


Figure 1.5: Total Income of the control group in the post DACA period (2014-2017). The estimated discontinuity is computed using linear regression; robust standard errors in parenthesis; the fitted values from the regression are plotted via lines.

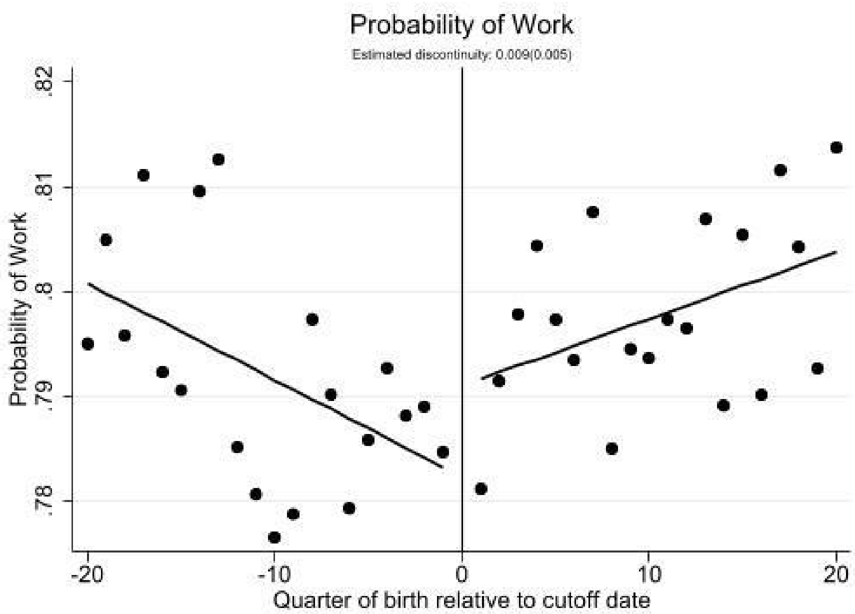


Figure 1.6: Probability of Work of the control group in the post DACA period (2014-2017). The estimated discontinuity is computed using linear regression; robust standard errors in parenthesis; the fitted values from the regression are plotted via lines.

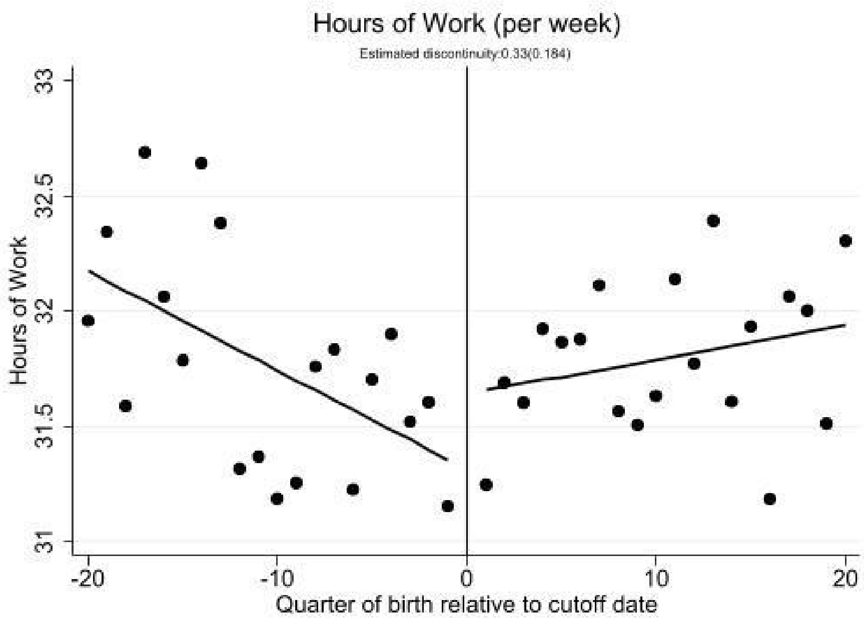


Figure 1.7: Hours of Work of the control group in the post DACA period (2014-2017). The estimated discontinuity is estimated using linear regression; robust standard errors are in parenthesis; the fitted values from this regression are plotted via lines.

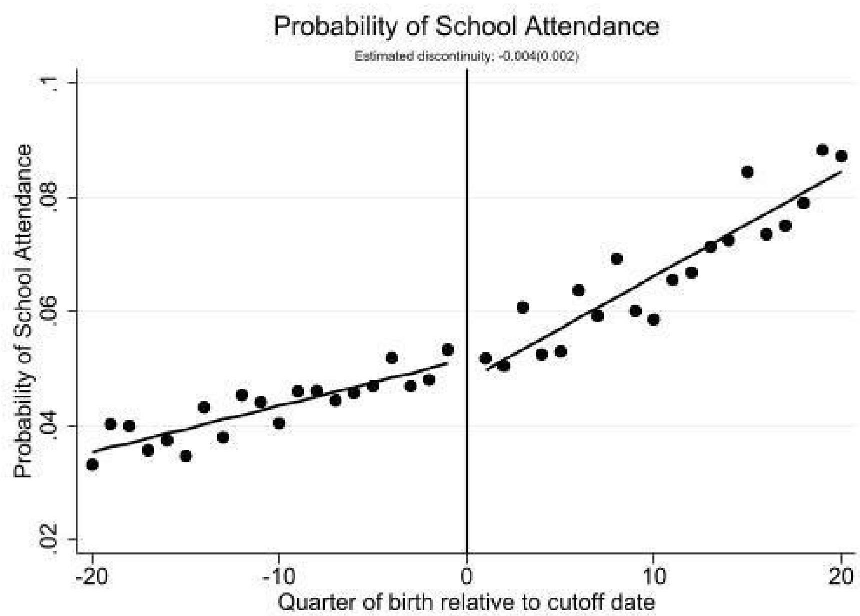


Figure 1.8: Probability of School Attendance of the treatment group in the post DACA period (2014-2017). The estimated discontinuity is computed using linear regression; robust standard errors in parenthesis; the fitted values from the regression are plotted via lines.

The results from the DRD analysis are depicted in Table 1.4. The results are robust in terms of their sign and significance to the estimates of the baseline RD estimates. There exist some differences in magnitude possibly due to differencing the impact that might have arisen due to the quarterly birth differences. Hence, the DRD supports the baseline RD analysis and implies towards increased income for the population satisfying the age criterion in the treatment group compared to the population in the control group satisfying the age criterion. The estimates range from $4,168 to $6,016, again implying improved financial stability.[14] The DRD analysis shows no evidence of increase in probability of work or current school attendance. [15]Furthermore, the results are indicative of an increase in hours of work.

The increased income implies a higher standard of living for the DACA eligible population. This could also be indicative of the DACA eligible population working for jobs that match their abilities (due to the gain of work authorization) rather than just applying for jobs that do not require paperwork or pay less due to the absence of paperwork.

My results are different from Pope (2016). While the results could arise from a difference in model specification, Pope (2016) also uses only two years of post DACA sample. Therefore, his study does not observe the large post DACA sample size that this study uses.

Table 1.4: DRD Results

Outcome Variable:	Total Income			Work or Not			Hours of Work			School Attendance		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Estimate	6016.271*** (2075.156)	5946.147*** (1827.840)	4167.918*** (1576.022)	0.024 (0.027)	0.029 (0.024)	0.022 (0.022)	2.129* (1.280)	2.302** (1.119)	1.444 (1.025)	0.015 (0.014)	0.009 (0.012)	0.010 (0.011)
Bandwidth	16	20	24	16	20	24	16	20	24	16	20	24
Mean	27,834.21	27,855.32	27,784.51	0.794	0.796	0.797	31.743	31.811	31.813	0.054	0.055	0.056
Observations	130,449	162,946	195,628	130,449	162,946	195,628	130,449	162,946	195,628	130,449	162,946	195,628

Note: All estimates are weighted using sampling weights; Robust standard errors in parentheses The analysis is performed using equation (2) ; * p < 0.1,** p < 0.05, *** p < 0.01

[14] This increase in financial stability could arise from various sources. In order to see the potential source of this increase, I replace the outcome of total income by wage income and investment income, respectively. The results mainly attribute this increase in financial stability to wage income. I show the results in appendix (A) Table A7.

[15] Although I see no evidence of a difference in current school attendance, I see that the DACA eligible population has a higher probability of having a college degree (4.6% to 5.4%). Results available on request.

1.6.4 Robustness Checks

1.6.4.1 Alternative Control Groups

This section provides robustness checks to the earlier results. In order to ensure that the results are robust to the choice of control groups, I compute the results based on different choices of control group. Therefore, I now conduct the DRD analysis using the same treatment group, but a different control group.

To reiterate, my main DRD analysis was performed using a control group, which comprised of Mexicans who were not eligible for DACA (based on other criteria other than the age criterion). I now create a control group where everyone who is not in the treatment group defined earlier forms the control group.[16] [17] This includes US citizens, foreign workers from different countries, etc.[18] Similar to the process done for the control group comprising of only the Mexican population, I now analyze the outcome variables graphically for the control group used as robustness check, that is the group that consists of all the population that does not form the treatment group. Figures 1.9-1.12 show the discontinuity arising in total income, probability of work, weekly hours of work, and probability of school attendance at a bandwidth of 20 quarters, respectively.[19] The estimated discontinuities are insignificant for income. I see a significant negligible discontinuity for probability of work at bandwidths of 20 and 24, a negative discontinuity for school attendance at the bandwidth of 20 and 24, and a positive discontinuity for hours of work at the bandwidth of 24.

[16] Note that this control group will now include some eligible population. For example, it will now include individuals from other countries who are not Mexicans, but still eligible for DACA However, I estimate that the proportion of these individuals should be very small compared to the sample size created with the alternative control group. For instance, a study by Brookings (Top 5 countries of origin of DACA immigrants brookings.edu) . finds that among the population from El Salvador (the country with the second largest population of DACA applicants), only 4% of the total population was DACA eligible.

[17] Appendix (A) tables A4-A6 show the comparison between the treatment group and the newly defined control group.

[18] I also perform two additional robustness checks where the control groups are more refined, that is, the control groups are a narrower version of the previous control groups. The sample used for these control groups include (1) Mexican non-citizens with a high school degree (appendix (A) tables A9-A11 show the comparison between the treatment group and this control group .DRD results are presented in appendix (A) table A12) (2) Mexican non-citizens with a high school degree that satisfy the year of immigration requirements.(Appendix(A) tables A13-A15 show the comparison between the treatment group and this control group DRD results are presented in appendix (A) table A16).

[19] The graphs for variables at alternative bandwidths of 16 and 24 are shown in the appendix (A) figures A38-A45.

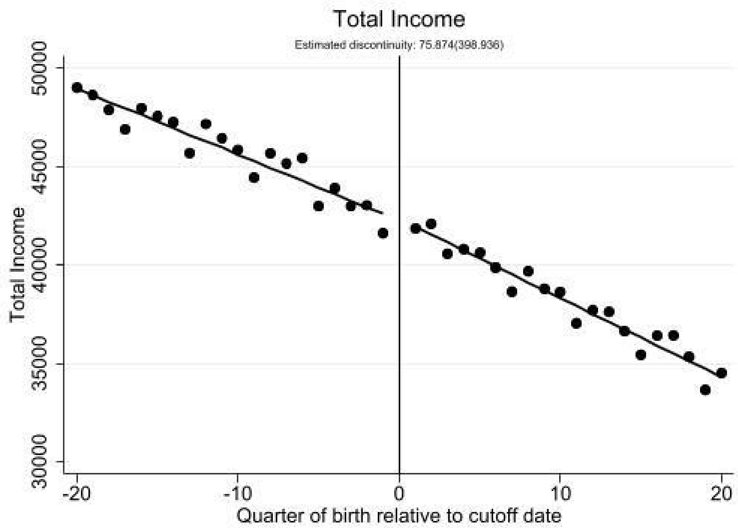


Figure 1.9: Total Income of the control group (used as a robustness check) in the post DACA period (2014-2017). The estimated discontinuity is computed using linear regression; robust standard errors in parenthesis; the fitted values from the regression are plotted via lines.

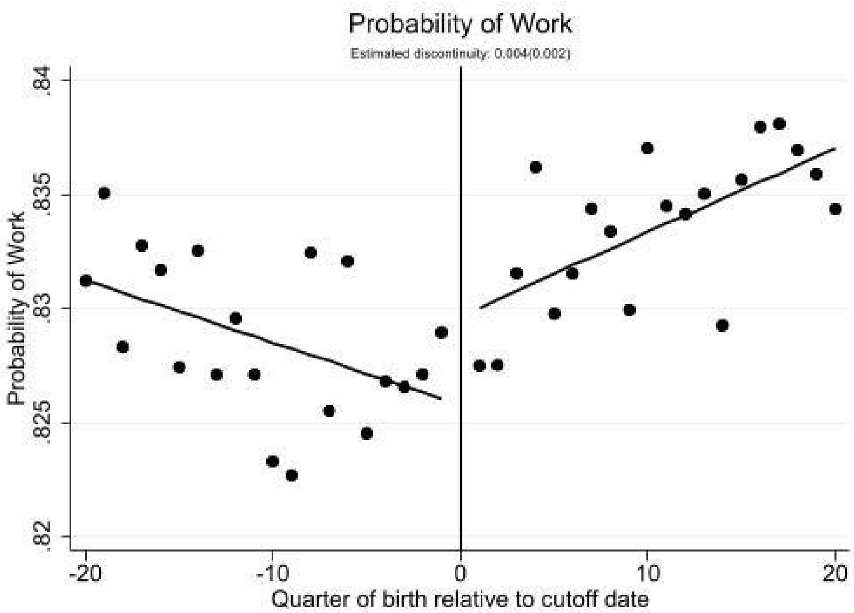


Figure 1.10: Probability of Work of the control group (used as a robustness check) in the post DACA period (2014-2017). The estimated discontinuity is computed using linear regression; robust standard errors in parenthesis; the fitted values from the regression are plotted via lines.

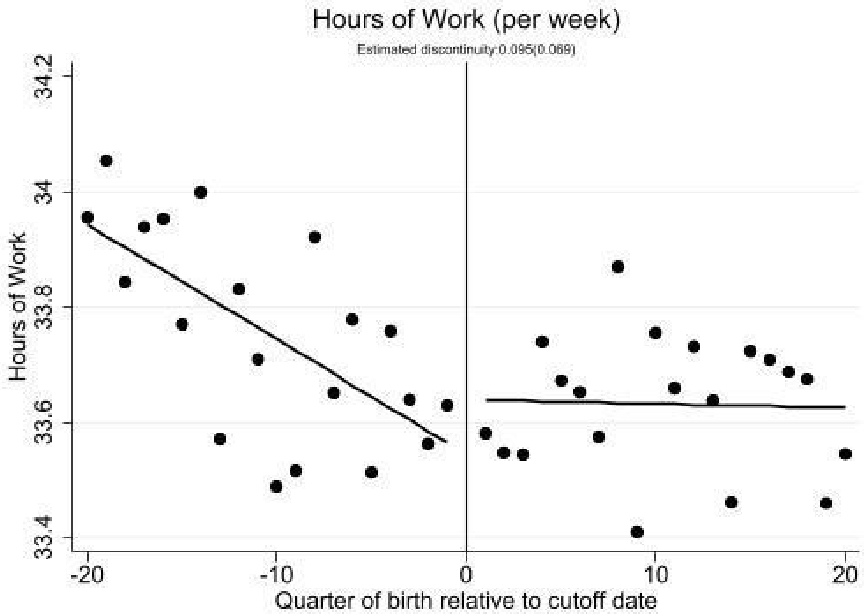


Figure 1.11: Hours of Work of the control group (used as a robustness check) in the post DACA period (2014-2017). The estimated discontinuity is estimated using linear regression; robust standard errors in parenthesis; the fitted values from this regression are plotted via lines.

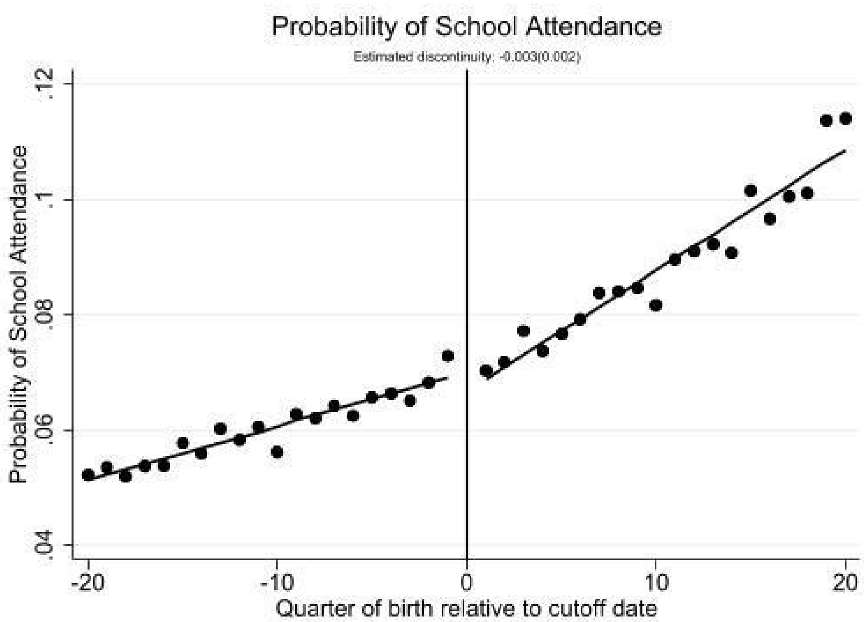


Figure 1.12: Hours of Work of the control group (used as a robustness check) in the post DACA period (2014-2017). The estimated discontinuity is estimated using linear regression; robust standard errors in parenthesis; the fitted values from this regression are plotted via lines.

The results from the DRD analysis using the alternative control group are depicted in Table 1.5. The results are robust in terms of sign and significance to the baseline RD analysis and the DRD analysis using the first control group. The results are again supportive of increased financial stability and labor force participation.

Table 1.5: DRD Results Using a Different Control Group

Outcome Variable:	Total Income			Work or Not			Hours of Work			School		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Estimate	6416.401***	6168.186***	4595.939***	0.030	0.033	0.027	2.444*	2.529**	1.644	0.014	0.009	0.012
	(2042.754)	(1797.997)	(1548.007)	(0.027)	(0.023)	(0.021)	(1.254)	(1.096)	(1.005)	(0.013)	(0.012)	(0.011)
Bandwidth	16	20	24	16	20	24	16	20	24	16	20	24
Mean	41,882.95	41,743.95	41,513.36	0.830	0.831	0.831	33.664	33.682	33.676	0.073	0.075	0.076
Observations	1,193,019	1,485,556	1,778,714	1,193,019	1,485,556	1,778,714	1,193,019	1,485,556	1,778,714	1,193,019	1,485,556	1,778,714

Note: All estimates are weighted using sampling weights; Robust standard errors in parentheses The analysis is performed using (2) . * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

1.6.4.2 Alternative Functional Form Specifications and Inclusion of Covariates in the Model

This section provides robustness checks to the DRD results derived from the original control group. I originally used a local linear specification. I now employ three additional functional form specifications: flexible linear, local quadratic, and flexible quadratic. Table 1.6 shows the results to these three alternative specifications and shows that the results stand robust. I also add fixed effects (for age, year of survey, gender, and state) and covariates (gender, disability variables) in the original model (Table 1.6). While my results are robust for the total income of the population, they lose significance for the hours of work.

Table 1.6: Robustness to Alternative Specifications and Inclusion of Covariates

Outcome Variable:	Total Income			Work or Not			Hours of Work			School Attendance		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Panel A: Flexible Linear Specification												
Estimate	5924.491***	5374.366***	3696.926**	0.022	0.025	0.022	2.156*	2.266**	1.486	0.014	0.007	0.007
	(2099.728)	(1840.708)	(1596.248)	(0.027)	(0.023)	(0.021)	(1.261)	(1.098)	(1.008)	(0.014)	(0.012)	(0.011)
Bandwidth	16	20	24	16	20	24	16	20	24	16	20	24
Mean	27,834.21	27,855.32	27,784.51	0.794	0.796	0.797	31.743	31.811	31.813	0.054	0.055	0.056
Observations	130,449	162,946	195,628	130,449	162,946	195,628	130,449	162,946	195,628	130,449	162,946	195,628
Panel B: Flexible Quadratic Specification												
Estimate	6097.248***	6132.557***	4263.740***	0.024	0.029	0.021	2.157*	2.342**	1.462	0.015	0.010	0.010
	(2111.031)	(1860.098)	(1603.095)	(0.027)	(0.024)	(0.022)	(1.279)	(1.120)	(1.027)	(0.014)	(0.012)	(0.011)
Bandwidth	16	20	24	16	20	24	16	20	24	16	20	24
Mean	27,834.21	27,855.32	27,784.51	0.794	0.796	0.797	31.743	31.811	31.813	0.054	0.055	0.056
Observations	130,449	162,946	195,628	130,449	162,946	195,628	130,449	162,946	195,628	130,449	162,946	195,628
Panel C: Local Quadratic Specification												
Estimate	5910.563	6241.085**	8378.242***	0.035	0.028	0.041	2.248	2.199	3.297**	0.032	0.031	0.021
	(3645.307)	(3059.049)	(2681.561)	(0.045)	(0.039)	(0.035)	(2.140)	(1.837)	(1.655)	(0.021)	(0.019)	(0.017)
Bandwidth	16	20	24	16	20	24	16	20	24	16	20	24
Mean	27,834.21	27,855.32	27,784.51	0.794	0.796	0.797	31.743	31.811	31.813	0.054	0.055	0.056
Observations	130,449	162,946	195,628	130,449	162,946	195,628	130,449	162,946	195,628	130,449	162,946	195,628
Inclusion of Fixed Effects (Survey Year, Age, Age Entered USA, and State Fixed Effects) and Covariates (Gender, Disability Variables)												
Estimate	5327.123***	4837.233***	3667.899**	0.003	0.002	0.007	1.193	1.020	0.568	0.017	0.011	0.012
	(2034.461)	(1781.970)	(1540.439)	(0.027)	(0.023)	(0.021)	(1.203)	(1.060)	(0.965)	(0.014)	(0.012)	(0.011)
Bandwidth	16	20	24	16	20	24	16	20	24	16	20	24
Mean	22,524.76	22,602.87	22,616.38	0.769	0.772	0.772	30.543	30.650	30.671	0.033	0.034	0.035
Observations	63,842	79,674	95,694	63,842	79,674	95,694	63,842	79,674	95,694	63,842	79,674	95,694

Note: All estimates are weighted using sampling weights; Robust standard errors in parentheses The analysis is performed using (2) . * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

1.7 Conclusion

This study studies the labor market impact of DACA. It finds evidence that DACA increases total income and hours of work of the potentially eligible DACA population. This is indicative of increased financial stability and labor force participation. It finds no evidence that DACA affects the likelihood of current school attendance of the potentially eligible DACA population.

The increased benefits of DACA point towards how the policy has improved the lives of unauthorized immigrants. Without the policy, these individuals live in a constant fear of deportation and are unable to make use of opportunities that their counterparts' avail. My study supports a potential pathway to permanent residency for the DACA population.

Chapter 2

Food Choices at a Client Choice Food Pantry: Do Low-Income Pantry Users Respond to Changed Opportunity Costs?

2.1 Introduction

53 million Americans received food at a food pantry in 2021. Over time, the quantity and quality of the food distributed through the food pantry system has improved, and client choice distribution models have been established as a best practice. In client choice models, individuals select food in settings similar to retail grocery environments, and there is less food waste compared to pantry models that provide a fixed assortment of foods. However, relatively little attention has been given to the ways in which the client choice pantry program design might impact the composition of individual food baskets that food pantry clients select.

In particular, there has been no discussion of how economic incentives might play a role in pantry food choices. It is not obvious that economic incentives should play a role because pantry foods are, by design, free. However, food choices in the pantry still represent trade-offs among the different foods available, and these trade-offs offer a possible role for economic incentives. Client choice food pantry designs differ in small ways that affect the way in which clients encounter these trade-offs. For example, some pantries place limits on the total quantity of food selected, limits on the amount that can be selected among individual items or groups of items, or some combination of different types of limits. Other pantries don't have explicit limits, but social norms, transportation constraints, and storage capacity create effective limits. Limits impose trade-offs related to food selections, and trade-offs mean that choosing one food item over another is costly. Economic incentives refer to strategies that modify the costs associated with choices, and they may play a role in client choice food pantries.

We examined a unique natural experiment in which a large client choice food pantry modified the choice architecture used within the client choice model. When the choice architecture changed, the set of opportunity costs associated with food choices at the pantry were modified. Our results suggest that individuals responded to the changed opportunity costs consistent with what economic theory would predict: foods in which opportunity costs declined were chosen more frequently. While the present study does not identify a new preferred approach to client choice food pantry programming, it does begin to establish an evidence base for the efficacy of incentives and choice architecture within the pantry context.

Crossroads Community Services (Crossroads) operates a client choice pantry located in Dallas, TX and is the largest non-profit food distributor in the 13-county North Texas region. Many food pantries give households similar quantities of food regardless of household size. However, while equitable in the sense that every client gets the same quantity and quality of food, this system does not account for varied household sizes. Crossroads pioneered an individualized system for allocating food distributions. Under this system, households were allocated points based on age, gender, and physical demands of employment of each household member. Households received enough points to acquire food from the pantry to provide 21-meals for each household member according to the United Stated Department of Agriculture (USDA) 2015-2022 dietary guidelines for Americans. Points functioned as a form of currency within the client-choice pantry.

When Crossroads began the points system, they focused on encouraging clients to select a balanced food basket. During this period, which we call the "Balanced Diet Regime," points were allocated in four categories (fruits/vegetables, grains, dairy and protein) in amounts that allowed each household to acquire the balance of food groups recommended by the USDA nutrition guidelines. The Balanced Diet Regime began on September 1, 2016 and ended on August 31, 2018. Individuals could select products within each category until their points for that category were exhausted.

On September 1, 2018, Crossroads abandoned the food category restrictions. Under the new "Open Regime", points were allocated as a lump sum and could be used to acquire food from as many or as few categories as desired.

During both regimes, the algorithm for allocating points based on household composition was unchanged. A 4-member household composed of mother, father, and two children aged 4-8 years received 144 points (150 points for the father, 136 points for the mother, and 79 points for each child) in both regimes. Also, the points required to acquire foods were held constant. For example, a 14.5 ounce can of green beans was 1 point and a pound of lean ground beef was 12 points in both the Balanced Diet and Open Regimes. Finally, during both regimes, limits were occasionally put in place on some products due to limited supply.

Table 2.1 shows the points needed to obtain foods that were available at the pantry on at least half of the days during the study window. The correspondence between points and the quantity of food is relatively consistent within categories but differs more significantly between categories. Proteins and Grains are expensive relatively to fruits, vegetables, and dairy products so the opportunity cost of selecting proteins and grains increased in the Open Regime compared to the Balanced Diet Regime. For example, in the Balanced Diet Regime, the opportunity cost of one pound of lean ground beef (12 points) may be 1 dozen eggs (12 points). In the Open Regime, the opportunity cost of one pound of lean ground beef (12 points) might also be 12 pounds of fresh carrots (12 points) or 12 cans of green beans (12 points). The information available in Table 1 for the full list of products available at the pantry during the study window is provided in Appendix Table B1.

As regimes changed, choice architecture changed because the structure of the decisions that clients had to make were altered. Clients were required to make more decisions in the Open Regime compared to the Balanced Diet Regime. On average during the study window, there were 34 different food options each day at the pantry: 6 food options each for fruits, grains, and protein foods; 2 food options for dairy products; and 14 food options for vegetables. In the Balanced Diet regime, decisions were made comparing options within a single category (at most 14 possible comparisons); and in the Open Regime, decisions were made comparing all 34 options. Under the Open Regime, clients were thus choosing both the point allocation between categories and the food products within a category, while under the Balanced Diet Regime only the later decisions were necessary.

2.2 Methods

Fixed effect panel data models were used to estimate within client changes in food choices (*Y*) for each individual *i* and period *t* as described in equation (1).

$$Y_{it} = \alpha + \beta_1 R_t + \beta_{2.} C_{it} + \beta_{3\,.} X_{it} + \eta_i + \varepsilon_{it} \qquad (1)$$

The Open Regime period is indicated by *R*; the vector C, was included to control for the number of unique food options available within each food category; X is a vector of client time-varying characteristics including monthly income of the household, SNAP amount received by the household, household size, and presence of kids in the household; η_i is the client-specific fixed effect.

In separate models, various outcomes were examined. These outcomes included the percentage of dietary points used for proteins, vegetables, dairy, grains, and fruits. Also, we examined the total number of different food categories from which selections were made during the pantry visit and the proportion of points used to obtain different types for produce (i.e., fresh, frozen, or shelf-stable fruits and vegetables). Finally, using a subsample of the data collected from visits in which household food security was assessed, we examined the moderating effect of food insecurity on response to the regime change.

2.3 Data

Analysis data was extracted from Crossroads administrative data and was collected from pantry clients during each of their food pantry visits from January 1, 2018 to December 21, 2018. Each observation in the data represents a visit to the pantry. Because clients may visit monthly, but not all choose to do so, the data is an unbalanced panel. Each observation in the data contains information on the date the household visited CCS, what food items were available and the points that each item "cost", the food items selected and all covariates included in our models. Crossroads administrative data has a high response rate, but not all variables were captured by every client at every visit. Sometimes clients refused to provide information and sometimes processes broke down at the pantry and data was not recorded. Nevertheless, we found that 95% of clients with records of food receipts had client data available for those visits. The analysis sample was restricted to those clients with at least one visit with complete data during each of the regimes.

The dependent variables used to characterize food choices were calculated as the percentage of dietary points spent by a household during a given visit within a particular food category. Food categories included the following: proteins, fruits, vegetables, dairy products, and grains. The percentage of points spent in a particular category was calculated as the number of points spent by the household in the category of interest divided by the sum of points used by the household across all categories during the same visit. We adopted the same procedure to compute the percentage of points spent on fresh, frozen, and shelf-stable produce by a given household at a particular visit.

2.4 Results

Table 2.2 provides summary statistics for the analysis data. Overall, 560 unique households visited during both regimes, and they were observed 4,069 times in total; on average, each household visited the pantry 7 times. The average SNAP amount received by households was $75 while the average monthly income was $1,094. All covariates had a higher between variation than within variation. For the outcome variables, on average, the largest proportion of points were spent on proteins (38%) and the least on dairy products (0.8%).

Summary statistics for the Balanced Diet and Open regime periods separately are available in the appendix (table B2 and B3). During the Open Regime, average point expenditures on proteins and average number of food categories selected was higher compared to the Balanced Diet Regime. The standard deviation of the proportion of points spent across different categories was higher in the Open Regime compared to the Balanced Diet Regime. Hence, we see an increase in the volatility in the proportion of points spent in the Open Regime.

2.4.1 Within-Household Changes

Estimation results for fixed effect models examining within-household changes in the proportion of points spent in each of the five food categories for households who participated in both the Balanced Diet and Open Regimes are reported in Table 2.3. The proportion of points allocated to dairy products declined when household size increased or kids were added to the household. Because household composition changes affect the total amount of points allocated to the household, these results do not necessarily indicate a change in the quantity of dairy products consumed by each household member. Other time-varying household characteristics had little impact on the allocation of points between categories. In general, when there were increases in the options available in a particular food category, the proportion of food points used in that same category increased, and the proportion of food points used in other food categories decreased.

The primary variable of interest was the indicator for observations during the Open Regime. Results suggest that there was a significant change in the way households allocated their dietary points between food categories when the policy changed from the Balanced Diet to the Open Regime. During the Open Regime, the percentage of points used for vegetables, fruits, and dairy products increased, and the proportion of points used to acquire grains and protein products decreased. The categories most impacted by the policy change were vegetables and grains; vegetable point allocations increased on average by 6.8 percentage points and grain point allocations decreased on average by 5.6 percentage points post policy.

Next, we examined how the policy change affected the total number of categories from which food selections were made—a measure of the variety of the food bundle obtained from the pantry. We also investigated the allocation of the change in proportion of vegetable and fruit points between fresh vegetables/fruits, frozen vegetables/fruits, or shelf stable types of vegetable/fruits. Household food bundles included a lower number of categories after the change to the Open

Regime (Table 2.4). Also, during the Open Regime, clients increased the proportion of points in all three types of produce, with larger increases for fresh and shelf-stable produce (Table 2.4).

Individuals who visit food pantries vary in their food needs. Some are highly food insecure and come to the food pantry during a time of crisis. Others have come to use the pantry as a resource to prevent crisis and thus are able to maintain a higher level of food security. Food security is assessed at Crossroads using the 10-item USDA food security module at every-other pantry visit. In Table 2.5, we present estimation results for our main models including an indicator for low or very low food security and assess the moderation effect of food security status on clients' response to the regime change. The models presented in Table 2.5 only include observations for pantry visits in which food security was assessed. Nevertheless, coefficient estimates are very similar to those reported in Table 2.3. Food security status does moderate the effect of the regime change in the case of vegetable and dairy point allocations. On average, after the regime change, food secure households increased their allocation of points in the vegetable category by 8.4 percentage points and households experiencing low or very low food security increased their vegetable point allocation by 5.5 percentage points. Similarly, point allocations in the dairy category increased by 0.7 percentage points among food secure households and 0.4 percentage points among food insecure households. Food security status did not significantly moderate changes in the protein, fruit, and grain categories.

2.5 Discussion

Our study highlights the potential for pricing incentives to impact client choices in the food pantry setting. Opportunity costs for grains and proteins increased when changing from the Balanced Diet Regime to the Open Regime, and the proportion of points used for grains and proteins also decreased when changing from the Balanced Diet Regime to the Open Regime. During the Open Regime, clients also shifted their selected food bundle to include proportionally more vegetables, fruits, and dairy products, which all became relatively less expensive under the Open Regime compared to the Balanced Diet Regime.

Importantly, these responses to price incentives occurred without promotions or any visible price changes. Client choices changed alongside changed opportunity costs or relative prices. A review of the literature on consumer response to opportunity costs found that resource constrained individuals may both have more incentive to consider opportunity costs (because tradeoffs are more salient) and be more likely to act impulsively (limited bandwidth or psychological responses to deprivation are among hypothesis for impulsivity). Thus, the extant evidence was unclear about whether food insecure populations would be likely to respond to opportunity costs. Our results suggest that food pantry clients were responsive to changed opportunity costs even when faced with severe resource constraints related to food.

The magnitude of the response to regime change is also non-trivial. The increase in vegetable point allocations is roughly equivalent to a 1-standard deviation increase in vegetable point allocations (standard deviation of within household vegetable point allocations was 6.8 percentage points (Table 2.2)). Changes in allocations in the dairy and grain categories represent nearly a half standard deviation change. Changes in allocations in the protein and fruit categories were comparatively smaller.

Substantial literature has studied client choice in the commercial grocery setting. Different techniques can be adopted to encourage clients to select certain food items including promotion, product placement, and pricing. A recent review found that food prices frequently affect product selection, and they are the preferred approach for impacting choices of low-income consumers.

Our results suggest that pricing incentives may be adapted to the food pantry setting by considering opportunity costs.

Limited work examining consumer choice in the pantry setting has focused on encouraging nutritious food selection through signage. The Supporting Wellness at Pantries (SWAP) program was designed in 2016 and modified in 2020 to begin to address the gap in nutrition promotion interventions developed for the food pantry setting. SWAP uses a color-coded rating system to identify and categorize foods as those that should be chosen often (green), sometimes (yellow) or rarely (orange). SWAP was evaluated in a sample of 222 households, and the nutritional quality of foods selected at the pantry after SWAP implementation improved modestly.[15] While SWAP focuses on providing point of purchase (or acquisition in the case of the food pantry setting) identification of nutritious food options, it does not attempt to use incentives to prioritize healthy choices.

In contrast, the change from Balanced Diet Regime to Open Regime that we studied may be viewed as a pricing strategy because the regime change impacted the opportunity cost of food selections. Prices for individual food items remained constant across both regimes, but opportunity costs changed when points were no longer restricted to being used within a single category. During the Balanced Diet Regime, points had to be used within a single food category, so cross-category "price" comparisons (e.g., the number of points needed for an additional can of vegetables versus an additional pound of ground beef) were irrelevant. In the Open Regime, pantry clients could use points in any food category making cross-category price comparisons an important decision input. Unfortunately, the natural experiment we investigated does not allow us to estimate price elasticities of demand to compare the impact of relative price changes in the food pantry setting with estimates for price elasticities in the retail setting.

The regime change undertaken by Crossroads was motivated by a desire to examine whether limiting client choices in the Balanced Diet Regime was necessary to achieve health promoting food selections at the pantry. Our results find no evidence that choice restrictions of the Balanced Diet Regime are necessary for health promotion at the pantry. However, this does not mean that pantry programming did not benefit from the Balanced Diet Regime. The Balanced Diet Regime began prior to 2010, a time when most food pantries were prioritizing serving a large number of households and moving a large amount of food with less regard for the nutritional quality, variety, or equity of food provided to households. The Balanced Diet Regime and the corresponding point system established to implement it institutionalized prioritization of procuring nutritious foods and a system for promoting equitable food distribution, which remain challenges for many pantries. However, in the current environment, benefits of the Balanced Diet Regime choice architecture are difficult to identify. To the extent that increasing fruit and vegetable consumption and reducing consumption of complex carbohydrates is health-improving, our results suggest that when constraints on food selections were reduced and opportunity costs changed, clients modified their behaviors in ways that may be health-improving. However, several caveats should be further examined before we conclude that the Open Regime resulted in beneficial behavior changes. First, not all foods in the fruit and vegetable categories are high in their nutritional content. For example, canned vegetables are often high in sodium, and fruit juices often have high amounts of added sugars. Likewise, not all foods in the grain and protein categories are nutritionally inferior.

When we tried to examine what types of produce selections increased, we found that the increase was distributed across fresh and shelf stable produce while frozen produce increased less. It is possible that there was only a small change in the amount of frozen produce selected because households are limited in their freezer capacity. Additional work should be undertaken to understand the nutritional implications for changes in food selection decisions and their relationship to food storage limitations and programmatic features.

Our results suggest a role for layering different approaches in designing optimal choice architecture for optimal pantry nutrition promotion. In nearly all models and across both regimes, when more options were available in a particular food category, the number of food items selected in that category increased. Thus, for example, to promote fresh fruit and vegetable consumption, pantries might prioritize offering a large number of fresh fruit and vegetable options relative to other food groups.

The most significant limitation of our study is the limited context that we were able to examine. Our data only allowed us to observe the food selections made at Crossroad's food pantry; we did not observe foods purchased at retail outlets or received from neighbors, family or friends, or food received through other programs such as the school lunch program. We are unable to determine whether the regime change resulted in changes in actual food consumption. Thus, any conclusions about the overall nutritional or economic impacts of the programmatic change can only be speculative. Additionally, our data comes from clients of a single, large food pantry and the external validity to other populations should be considered carefully.

Another potential limitation in our model is our inability to control for Crossroads food procurement process, which might lead to a simultaneity problem. For instance, in the open regime, when the households were not constrained by limits and could choose items freely, some food items might have had a higher demand. This might create a concern that the food pantry would supply more of those goods to make sure that the households are able to acquire what they want thus changing food inventories. However, we can temper the concern by noting that (1) Crossroads has limited ability to affect pantry supply because they are limited to the supply available from the regional food bank and typically, under both regimes, Crossroads received as much food and as much variety as possible—thus there was almost no margin over which to make changes and (2) the food bank does in some cases adjust their food procurement strategy in response to pantry feedback, but this is a long-term process. Our study window only included 4 months in the Open Regime, which is too short of a time for Crossroads to provide feedback to the food pantry regarding supply preferences and then have changes materialize at the pantry.

2.6 Conclusion

Our study provides evidence that choice architecture influences food selections among pantry clients. The number of food options available in a particular category and the opportunity costs of choosing different foods from a client choice pantry influenced food selections. Even though pantry food is "free", opportunity costs still exist and provide economic incentives that impact food choices. Future work aimed at promoting nutrition among food pantry clients may benefit from incorporating economic incentives through pantry choice architecture.

Table 2.1: Point Values Assigned to Food Items at the Pantry

Protein Foods

Food Item	Size	Points Required for Purchase	% of days offered during study window
Beef Stew - USDA	24 oz bag	24	58.79
Eggs - Whole Fresh - Small	1 doz	12	50.75
Peanut Butter (Algood)	18oz jar	18	50.25

Dairy Products

Food Item	Size	Points Required for Purchase	% of days offered during study window
Dairy - 1% Natrel Shelf Stable	32 oz	1	62.31

Vegetables

Food Item	Size	Points Required for Purchase	% of days offered during study window
Green Beans - Signature	14.5 can	1	74.87
Corn (Signature)	15.25 oz can	1	73.37
Beans Dry - Dark Red Kidney Beans	2 lb bag	6	69.35
Onions - Fresh	4lb Portion	4	65.33
Pinto Beans - Dry - 32 oz bag	2 lb bag	6	55.78
Carrots - fresh	4 lb portion	4	52.26

Grains

Food Item	Size	Points Required for Purchase	% of days offered during study window
Spaghetti (M)	16oz box	16	78.89
Cereal - Bran Flakes (RF)	17.3 oz box	17	60.80
Oatmeal - Quick Cooking Rolled Oats	3 lb bag	48	51.76

Fruits

Food Item	Points Required for Purchase	Points Required for Purchase	% of days offered during study window
Blueberry -frozen 3 Lb bag	3lb bag	6	74.37
Figs - Dry	1 lb bags	3	65.83
Plums - Dry Pitted	1 lb bags	3	54.77

This table only includes those food items offered on 50% or more of the days during the study window. Please see Table A1 in the appendix for a complete listing of all food items offered.

Table 2.2: Summary Statistics for Covariates and Outcomes for Clients who Experienced Both Regimes (N=560 households and 4069 observations*)

Variables		**Mean**	**Std. Dev.**	**Min**	**Max**
Covariates					
Household Income ($)	Overall	1094.025	782.018	0	5640
	Between		654.302		
	Within		464.126		
SNAP Amount	Overall	75.380	140.283	0	794
	Between		139.345		
	Within		50.616		
Household Size	Overall	4.475	2.585	1	13
	Between		2.538		
	Within		0.396		
Indicator of Kids in Household	Overall	0.640	0.480	0	1
	Between		0.478		
	Within		0.073		
Outcome Variables					
Percent- Protein	Overall	0.379	0.138	0	1
	Between		0.070		
	Within		0.125		
Percent-Vegetables	Overall	0.146	0.074	0	0.718
	Between		0.037		
	Within		0.068		
Percent- Dairy Products	Overall	0.008	0.013	0	0.303
	Between		0.007		
	Within		0.012		
Percent- Grains	Overall	0.401	0.149	0	1
	Between		0.068		
	Within		0.137		
Percent-Fruits	Overall	0.064	0.040	0	0.545
	Between		0.021		
	Within		0.035		
Percent-Miscellaneous	Overall	0.001	0.003	0	0.026
	Between		0.002		
	Within		0.001		
Number of Categories	Overall	4.794	0.826	1	6
	Between		0.662		
	Within		0.526		
Percent-Fresh Produce	Overall	0.104	0.049	0	0.397
	Between		0.026		
	Within		0.043		
Percent-Frozen Produce	Overall	0.010	0.017	0	0.250
	Between		0.009		
	Within		0.015		
Percent-Shelf Stable Produce	Overall	0.096	0.066	0	0.636
	Between		0.033		
	Within		0.060		

*On average, each household visited 7.27 times.

Table 2.3: Estimated Coefficients for Models Examining Point Allocations Across Food Categories

VARIABLES	(1) Proteins	(2) Vegetables	(3) Fruits	(4) Grains	(5) Dairy Products
Open Regime Indicator	-0.025***	0.068***	0.008***	-0.056***	0.005***
	(0.006)	(0.003)	(0.002)	(0.005)	(0.001)
Household Income	0.000***	-0.000***	-0.000*	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
SNAP Amount	0.000	0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Household Size	-0.002	0.004	0.001	-0.002	-0.001***
	(0.004)	(0.003)	(0.001)	(0.003)	(0.000)
Indicator of Kids in Household	-0.011	-0.014	-0.000	0.030*	-0.004**
	(0.019)	(0.013)	(0.004)	(0.016)	(0.001)
[1]Options Available- Fruits	0.002	-0.011***	0.008***	0.001	0.000**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
[1]Options Available- Vegetables	-0.001*	0.005***	-0.003***	-0.001	0.000***
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)
[1]Options Available- Grains	-0.023***	-0.006***	-0.001***	0.030***	0.000
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)
[1]Options Available- Dairy Products	0.002	0.002	0.002**	-0.017***	0.011***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.000)
[1]Options Available- Protein	0.030***	-0.014***	-0.005***	-0.010***	-0.000***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)
[1]Options Available- Miscellaneous	0.016	-0.002	-0.004	-0.015	0.001
	(0.017)	(0.008)	(0.004)	(0.014)	(0.001)
Constant	0.361***	0.244***	0.077***	0.321***	-0.005***
	(0.031)	(0.019)	(0.008)	(0.026)	(0.001)
Observations	4,069	4,069	4,069	4,069	4,069
R-squared	0.422	0.441	0.340	0.594	0.607
Number of Unique Households	560	560	560	560	560

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

[1]Options Available measures the number of food options available to households in a particular food category at a given visit.
Dependent Variable times 100 is the percentage of points spent by households on that particular category at a given visit.
All models are estimated as individual fixed effect models of within-household changes.

Table 2.4: Estimated Coefficients for Models Examining the Food Bundle Variety and Types of Produce Selected

VARIABLES	(1) Food Bundle Variety (Number of Categories in Bundle)	(2) Fresh Produce	(3) Frozen Produce	(4) Shelf-Stable Produce
Open Regime Indicator	-0.334***	0.036***	0.002***	0.041***
	(0.027)	(0.002)	(0.001)	(0.003)
Household Income	0.000	-0.000	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
SNAP Amount	0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Household Size	0.044**	-0.001	0.002***	0.005*
	(0.020)	(0.001)	(0.000)	(0.003)
Indicator of Kids in Household	0.337**	-0.011	-0.001	0.002
	(0.150)	(0.009)	(0.002)	(0.010)
[1]Options Available- Fresh Produce		0.008***	-0.000**	-0.004***
		(0.000)	(0.000)	(0.000)
[1]Options Available- Frozen Produce		-0.003**	0.012***	-0.000
		(0.001)	(0.001)	(0.002)
[1]Options Available- Shelf-Stable Produce		-0.006***	-0.001***	0.002***
		(0.001)	(0.000)	(0.001)
[1]Options Available- Grains	0.023***	-0.003***	-0.000***	-0.005***
	(0.004)	(0.001)	(0.000)	(0.001)
[1]Options Available- Dairy Products	0.303***	0.000	0.001***	0.003**
	(0.011)	(0.001)	(0.000)	(0.001)
[1]Options Available- Protein	0.015***	-0.007***	-0.001***	-0.010***
	(0.005)	(0.000)	(0.000)	(0.001)
[1]Options Available- Miscellaneous	0.895***	-0.008	-0.001	0.003
	(0.046)	(0.006)	(0.002)	(0.006)
[1]Options Available - Fruits	0.018***			
	(0.006)			
[1]Options Available – Vegetables	-0.028***			
	(0.003)			
Constant	3.981***	0.163***	0.007**	0.162***
	(0.148)	(0.012)	(0.003)	(0.017)
Observations	4,069	4,069	4,069	4,069
R-squared	0.367	0.291	0.410	0.470
Number of Unique Households	560	560	560	560

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

[1]Options Available- Category measures the number of food options available to households in a particular food category at a given visit.

Dependent Variable (in column (1)) shows the number of categories a household chose from during a given visit.

Dependent Variable (in columns (2)-(4)) times 100 shows the percentage of points spent by households on that particular category at a given visit.

All models are estimated as individual fixed effect models of within-household changes.

Table 2.5: Moderation Effect of Food Insecurity

VARIABLES	(1) Proteins	(2) Vegetables	(3) Fruits	(4) Grains	(5) Dairy Products
Open Regime Indicator	-0.030**	0.084***	0.007**	-0.068***	0.007***
	(0.012)	(0.006)	(0.003)	(0.011)	(0.002)
Indicator for Food Insecurity* Open Regime Indicator	0.020	-0.029***	0.005	0.007	-0.003*
	(0.016)	(0.009)	(0.005)	(0.014)	(0.002)
Household Income	-0.000	-0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Indicator for Food Insecurity	-0.027***	0.011**	0.001	0.015*	0.001
	(0.009)	(0.005)	(0.003)	(0.009)	(0.001)
SNAP Amount	-0.000	-0.000	-0.000	0.000	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Household Size	0.005	0.002	-0.000	-0.006	-0.001**
	(0.006)	(0.003)	(0.002)	(0.007)	(0.000)
Indicator of Kids in Household	-0.039	0.011	-0.003	0.033	-0.002
	(0.028)	(0.023)	(0.006)	(0.028)	(0.002)
[1]Options Available- Fruits	0.003	-0.012***	0.009***	-0.001	0.000
	(0.003)	(0.002)	(0.001)	(0.002)	(0.000)
[1]Options Available- Vegetables	0.001	0.004***	-0.003***	-0.003**	0.000***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)
[1]Options Available- Grains	-0.021***	-0.005***	-0.001	0.027***	0.000
	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
[1]Options Available- Dairy Products	0.001	0.001	0.004***	-0.018***	0.012***
	(0.004)	(0.003)	(0.001)	(0.004)	(0.001)
[1]Options Available- Protein	0.026***	-0.014***	-0.004***	-0.008***	-0.000***
	(0.002)	(0.001)	(0.001)	(0.002)	(0.000)
[1]Options Available- Miscellaneous	0.013	-0.003	0.000	-0.016	0.002
	(0.026)	(0.015)	(0.006)	(0.023)	(0.001)
Constant	0.342***	0.224***	0.068***	0.371***	-0.007**
	(0.047)	(0.028)	(0.014)	(0.045)	(0.003)
Observations	1,465	1,465	1,465	1,465	1,465
R-squared	0.419	0.457	0.351	0.612	0.534
Number of Unique Households	524	524	524	524	524

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1
[1]Options Available- Category measures the number of food options available to households in a particular food category at a given visit.
Dependent Variable times 100 shows the percentage of points spent by households on that particular category at a given visit.
All models are estimated as individual fixed effect models of within-household changes.

Chapter 3

Do Long Term Cash Transfers Improve the Status of Girls?

3.1 Introduction

There exists a strong son preference in developing countries. A son is considered as an asset for the household, someone who is going to be the future breadwinner of the family and can take care of his parents. In contrast, a daughter is considered a liability due to being sent off to another household post marriage (specifically in the South Asian culture). This strong son preference has led to great sex selection in those countries. For example, female feticide[20], the abortion of a female fetus outside of legal methods, is prevalent.

India is one of the most noteworthy countries for sex selection. The worsening sex selection can be seen from the continuous deterioration of its sex ratio. The expected sex ratio at birth (male to female) is 1.05.[21] Whereas in India, for the age group of 0 to 6 years, the sex ratio was 102.4 males per 100 2001,s in 1961, 104.2 in 1980, 107.5 in 2001 , and 108.1 in 2011.[22] [23]For the age group of

[20] In 1994, the Parliament of India passed the Pre-Conception and Pre-Natal Diagnostic Techniques (PCPNDT) Act. The act banned the predetermination of the sex of the fetus and was aimed at putting a stop to the practice of sex determination before the birth of the child. However, there has been a concern that the act was not properly enforced, and that the practice of prenatal sex determination is still prevalent. (source:UNICEF India". UNICEF. Archived from the original on 2014-12-23. Retrieved 2012-05-06)
[21] https://ibis.health.utah.gov/ibisph-view/indicator/important_facts/SexRatBrth.html
[22] Data Highlights - 2001 Census Census Bureau, Government of India
[23] India at Glance - Population Census 2011 - Final Census of India, Government of India (2013)

0-1 years, India maintained a sex ratio of 109.9 with some states maintaining a sex ratio as high as 128.[24] [25]

In order to improve the situation of the girl child in the society, Bihar, a state in India, launched a policy called "Mukhyamantri Kanya Suraksha Yojana (MKSY)". MKSY had clear objectives. It aimed to improve the sex ratio, prevent female feticide, encourage birth registration, promote the birth of a girl child, and ensure her rightful place of pride in the society (Sekhar 2010). The policy promised to provide a long-term cash transfer if the girl child and the family satisfied a few eligibility conditions pertaining to household income, date of birth, birth registration, and birth order of the girl child.

In this study, I intend to study two potential impacts of the policy.[26] Firstly, I aim to see if there is an improvement in the survival rate of the girl child. Secondly, I intend to study if there was any difference in the schooling outcomes.

Using data from the National Family Health Survey (NFHS) and leveraging the local variation arising from the policy eligibility conditions, I employ a regression discontinuity design. I exploit the eligibility condition of the date of birth of the girl child and create the eligible and the ineligible groups to study any differences in outcomes. Subsequently, I compare the outcomes of those just eligible vs. those just ineligibles for the transfer based on whether their date of birth fell after or before the November 22, 2007, cutoff.

[24] Age Data - Single Year Age Data - C13 Table (India/States/UTs) Population Enumeration Data (Final Population) - 2011, Census of India, Ministry of Home Affairs, Government of India

[25] Age Data - Single Year Age Data - C13 Table (India/States/UTs) Population Enumeration Data (Final Population) - 2011, Census of India, Ministry of Home Affairs, Government of India

[26] Due to the policy initiation date in 2008, very limited outcome variables can be studied. This is because the girls at the time of the survey are not old enough to get married. Hence, their fertility outcomes, etc. cannot be studied.

I find no significant effects of the policy on the survival rate of the first-born girl child. However, I find some evidence of a higher probability of survival for the second born girl child where the first-born child is a female. This result is robust across alternative specifications. Furthermore, the results indicate that while such policies may help already enrolled first born children stay longer in school, they do not affect the probability of school attendance per se. My results for schooling outcomes are consistent with other studies (Sinha and Yoong 2009) studying the impact of similar long-term financial incentives in India.

This study contributes to the literature by assessing the efficacy of a policy specifically aimed at curbing female feticide. There exists little evidence in South Asia due to the recency of these policies (Sekher 2015). To the best of my knowledge, there exist four studies, which conduct an impact evaluation of policies launched to improve the sex ratio in India. Firstly, Sinha and Yoong (2009) conduct a formal impact evaluation of the Apni Beti Apna Dhan (ABAD) scheme. The policy provides an immediate cash grant to the eligible mother and a long-term savings bond to the eligible girl child. Furthermore, there exist additional cash grants based on the completion of education levels by the girl child. The study finds positive impacts on female child survival and education outcomes (Kumar and Sinha 2018). The second study is that of Sekher and Ram (2015) who study the Dhanlaxmi scheme. Although the study finds some immediate positive gains, there was no observed change in preference for the girl child. Additionally, Anukriti (2018) studies the impact of the Devirupak policy. The policy was initiated in the state of Haryana to lower fertility and improve the sex ratio. The results from the study find that the policy failed to change son preference. Finally, Jain (2022) studies the impact of the Ladli Laxmi Yojana (LLY), a scheme launched in Madhya Pradesh in the year 2006. The objective of the scheme was to encourage the birth of girls. The study finds that the policy increased the probability of a female birth with one son by around 18 percent.

This study contributes to the scant literature on conditional cash transfers (CCTs improving the sex ratio) in India by studying another policy, which gives the promise of a long-term incentive. Sinha and Yoong (2009) is the only study that studies the benefits of a long-term incentive, which promises a future gain. However, unlike the policy studied by them (ABAD), MKSY is unique in a sense that it offers no direct transfer/interim payments and is entirely a promise of future pay. The only benefit associated with the policy is the cash transfer paid to the girl child once she turns 18. Hence, I contribute to the literature by studying a CCT offering a long-term incentive in a different state (Bihar) compared to Sinha and Yoong (2009) who study a long-term policy (with interim payments) in the state of Haryana.

The rest of the paper progresses as follows. Section 2 describes the policy. Section 3 describes the data and the empirical strategy. Section 4 discusses the results. Section 5 concludes the study.

3.2 Institutional Background [27]

Mukhyamantri Kanya Suraksha Yojana was initiated in July 2008 by the Social Welfare Department, the Women Development Corporation (WDC) and the UTI asset management company. It estimated an amount of Rs 140 (approx. $17 million) to empower over 7 lakh girls and women in need. To popularize the scheme, Bhojpuri theatre and radio advertisements were used.

[27] The information regarding institutional design is taken from Sekhar(2010) .All the information included in his report (for Indian government and UNFPA) was collected during his visits to the states in 2010 as well as his interaction with senior state govt. officials responsible for managing the program.

In order to be eligible, the girl child in Bihar should belong to a below poverty line (BPL) family (when the child is born). Furthermore, only the first two girl children of the family are eligible for the policy. In addition, the girl child's birth must have been registered within a year, she should be born after Nov 22, 2007, and should be between the age of 0-3 years at the time of applying.

The policy invests Rs 2,000 for the girl child if the girl child and her family satisfy these eligibility conditions. When the girl child turns 18 years of age, she will be paid an amount equal to the maturity value. The maturity value is around Rs. 18,000 (approx. $218). In the case that the girl child dies during the period, the amount will be paid back to the Women Development Corporation.

In order to improve the status of the girl child, Bihar has launched several polices. These policies range from providing bicycles to girls to improve schooling outcomes to providing help for marriage expenses. These existing policies could have increased the uptake of the policy.

3.3 Data and Empirical Strategy

The data for this study comes from the National Family Health Survey (NFHS) conducted during the year 2015-2016.[28] NFHS interviews women aged 15-49 years. The data is one of the most reliable and popular sources of information that has been used in past studies (Sinha and Yoong 2009; Anukriti 2018). Additionally, the data, to the best of my knowledge, gives the best source of information that recognizes the conditions for policy eligibility (eg: birth order, BPL status of the household, date of birth of the child, etc.). The information for the outcome variables studied is found under different data sets and hence, needs to be merged.

[28] The data is available through the DHS website and is available upon request.

I measure two broad categories of outcomes in this study. The first category measures any differences in the probability of survival of the girl child (at the time of interview) eligible for the policy and the girl child ineligible for the policy. NFHS asks the question to the mother regarding whether the child of the mother is still alive (at the time of interview) or not. I leverage this information in the dataset. The second category measures the schooling outcomes of the girl child. These are measured by asking whether the child attended school during the current school year, whether the child ever attended school, education in years during the current school year, etc.

For studying whether the policy affected the survival rate of the girl child, I use the question in the Births Recode (BR) that asks a mother regarding whether her child was alive at the time of interview. Approximately 99.59% birth records had complete birth date information, that is, they did not have any imputation. Furthermore, I merge the BR file with the Households Recode (HR) file in order to match the births in the BR file to their household status, that is, whether the household belonged to a BPL category or not.

For the outcome measuring the education status, I merge the BR file to the Persons Recode (PR file). The PR file contains information about the schooling of members who stayed in the household last night and the usual residents of the household.

Due to the general low status of women in the state, Bihar has recently launched many policies for the empowerment of women. These policies range from providing free school uniforms to providing monetary help for a daughter’s marriage. The simultaneous existence of these policies makes a cleaner identification of the effect of a single policy (like the one currently being studied) cumbersome. Hence, assumptions of methods like difference in differences are hard to satisfy. Therefore, in order to estimate the effect of the policy, I employ a regression discontinuity approach by exploiting the date of birth eligibility condition for the girl child. Assuming that the date of the birth of the girl child is exogenous to the survival and education outcomes, the computed estimate at the cutoff will give the closest estimate to the true value.

For my empirical analysis, I create a sample that is eligible for the policy in every respect except for the date of birth of the girl child. Firstly, since the policy was initiated in Bihar, I create a sample of women who were interviewed in Bihar and exclude women interviewed in other states.

Secondly, following the eligibility condition, I only include women who belong to the BPL households.[29] This creates a sample that is comparable in every (observable) respect except the date of birth of the girl child.

For my study, I conduct the analysis for the first-born and the second born girl child of women.[30] Studies have proven that the sex of the first born is random in India and sex-selectivity starts at higher order births. However, my study intends to study investment on the first born. Hence, although there might not exist any sex selection on the first-born child, there might exist discrimination in terms of the resources provided to first born girls. For instance, according to (Jayachandran and Pande 2017) the birth of a daughter with no older brothers will cause the parents to reduce the resources spent on the daughter due to the possibility of exceeding their intended fertility in order to try again for a son. They further find that the first-born son enjoys a height advantage. However, there exists no such advantage for the first-born daughters in India. For the second-born girl child, I do two separate analyses. Firstly, I consider the second born girl child where the first-born child is a son. Subsequently, I consider the second born girl child where the first born is a daughter. I conduct separate analysis because any effects on the second girl child

[29] There has been a debate regarding whether the poorest population in India who should be holding a BPL card actually holds one. However, since the condition of availing the benefit is based on whether the household has a BPL card, I will use this measure.

[30] While I can take into account the heterogenous effects arising from the older sibling (for the second born girl child), my study cannot consider heterogenous effects from the children born at third order onwards, that is siblings born at later dates than the sample considered. This is because restricting to women with only two total births do not give me a large enough sample size; Sinha & Yoong (2009), to the best of my knowledge and interpretation, do not do this as well.

might be determined by the gender of the previous children born. Hence, in order to avoid any heterogenous effects arising from previous births, I consider two separate analyses for the second born child based on the gender of the first born.

The policy requires the girl child to be a permanent resident of Bihar.[31] Since there is no way to find out their true place of residence, following (Anukriti 2018), for both the categories of outcomes, I drop women who were not de jure residents (usual residents of the household). Additionally, I also drop women who had their first birth before the age of 10 years.

The potentially eligible group consists of girls who satisfy all the observable conditions of eligibility while the ineligible group consists of girls who satisfy all the observable conditions except the date of birth. I call the eligible group "potentially eligible" since I do not observe all the conditions of eligibility (whether the girl child is a permanent resident). I estimate the impact of being born after the policy cutoff date and hence being potentially eligible for the policy [32] , and therefore estimate the reduced-form effect of the policy instrument on outcomes, where the policy instrument is defined as the binary indicator of born before or after the cutoff date.

The running variable in the analysis is the month and year of birth of the child (since day of birth is not observed in the data)[33] . Births that take place before the eligibility date of November, 2007, are ineligible children while those born after the date of November, 2007, are potentially eligible children. Since day of birth is not observed, I exclude observations pertaining to November 2007 from the analysis as I cannot observe the exact eligibility for those born in November.

[31] In order to be a permanent resident, the person should reside in Bihar for three years, own land/plot in Bihar, and be on the country's voter's list. For minors, the eligibility is determined based on parent's eligibility.
[32] I cannot condition on birth registration status as it is an outcome variable that changes at the cutoff itself due to this policy).
[33] This is not a problem in general, as long as one interprets the estimated parameter accordingly as explained in (2021, https://arxiv.org/abs/2111.07388).

Equation (1) denotes the reduced form model.

$$Y_i = \delta_0 + \delta_1 R_i + \delta_2 D_i + \delta_3 R_i D_i + \epsilon_i \quad (1)$$

where R_i refers to the running variable, that is, the normalized (recentered) year and month of birth of an individual with November 2007 as the cutoff ; D_i indicates whether the girl child was eligible for the policy. I use robust standard errors. Furthermore, I use women's (state) sampling weights due to the stratified nature of data collection.

For measuring the difference in survival rate of the girl child, I consider a maximum bandwidth of three years around the cutoff. I perform the analysis at 6 months, 18 months, and 36 months. For studying the schooling outcomes, I conduct the analysis for two different bandwidths that span 12 months before and after cutoff due to the school age eligibility criteria of being 6 years of age in April.

δ_2 is the parameter of interest, which estimates the reduced-form effect of the policy instrument on outcomes, where the policy instrument is defined as the binary indicator of born before or after the cutoff date.

Table 3.1 provides the descriptive statistics of the sample used for the first outcome variable (survival rate). Table 3.2 provides the descriptive statistics of the of the sample used for the second outcome variable (education outcomes).[34]

[34] One of the potential weaknesses of the data for the education outcomes is the small sample size.

Table 3.1: Descriptive Statistics (Sample for Outcome 1: Whether the Child is Alive at the Time of Interview)

Variable	Mean		Difference	t-statistic
	Above Cutoff	Below Cutoff		
Alive	0.955	0.946	0.009	1.20
Poor	0.869	0.856	0.013	1.11
Middle Income	0.088	0.104	-0.016	-1.55
Rich	0.043	0.040	0.003	0.39
Hindu	0.845	0.843	0.002	0.14
Muslim	0.153	0.156	-0.003	-0.20
Urban	0.067	0.078	-0.011	-1.26
Education of Woman	2.618	1.981	0.637	4.79
Observations	2,030	2,080		

Binary variables are coded in percentage terms; Observations “above cutoff” refer to the sample that is born after Nov 22, 2007 (hence, policy eligible based on the age discontinuity condition) and observations “below cutoff” refer to the sample that is born before the cutoff of Nov 22, 2007 (hence, policy ineligible based on the age discontinuity condition); women’s state weights used

Table 3.2: Descriptive Statistics (Sample for Outcome 2: Schooling Outcomes)

Variable	Mean		Difference	t-statistic
	Above Cutoff	Below Cutoff		
Attended School During Current School Year	0.840	0.907	-0.067	-3.31
Ever Attended School	0.845	0.910	-0.065	-3.23
Grade of Education During Current School Year	1.480	2.135	-0.655	-9.05
Education in Years (Current School Year)	1.484	2.168	-0.684	-8.48
Poor	0.859	0.832	0.027	1.21
Middle Income	1.094	0.128	-0.034	-1.69
Rich	0.045	0.039	0.006	0.52
Hindu	0.833	0.844	-0.011	-0.46
Muslim	0.164	0.156	0.008	0.33
Urban	0.055	0.085	-0.030	-1.96
Education of Woman	2.507	2.033	0.474	2.04
Observations	645	680		

Binary variables are coded in percentage terms; Observations “above cutoff” refer to the sample that is born after Nov 22, 2007 (hence, policy eligible based on the age discontinuity condition) and observations “below cutoff” refer to the sample that is born before the cutoff of Nov 22, 2007 (hence, policy ineligible based on the age discontinuity condition); women’s state weights used

3.4 Results

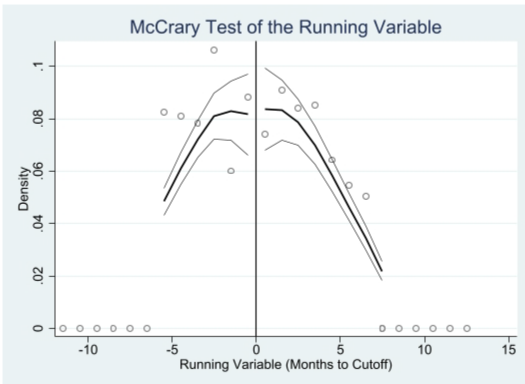
3.4.1 Validity Checks[35]

I first begin by testing the assumptions required for a regression discontinuity design for the combined sample of the first born and the second born girl child. I begin by checking the assumption of manipulation, that is, I test for any sorting of individuals around the threshold. Figure 3.1 shows the McCrary test at 6 months, 18 months, and 36 months for the sample created for the first outcome variable. Table 3.3 shows the statistical tests. I witness no evidence of manipulation. Similarly, Figure 3.2 and Table 3.4 show the same test for the sample created for the education outcome variables at 6 months and 12 months. I again witness no evidence of manipulation.

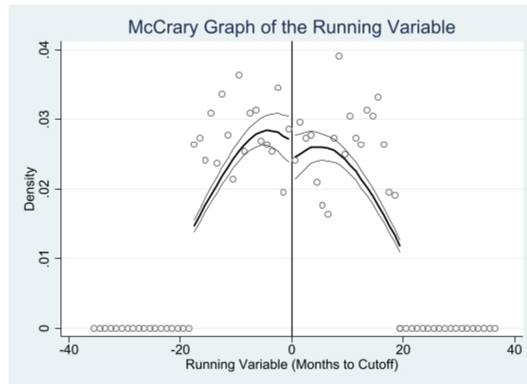
[35] The validity check is done for the combined sample of the first born and the second born girl child.

McCrary Tests[36]

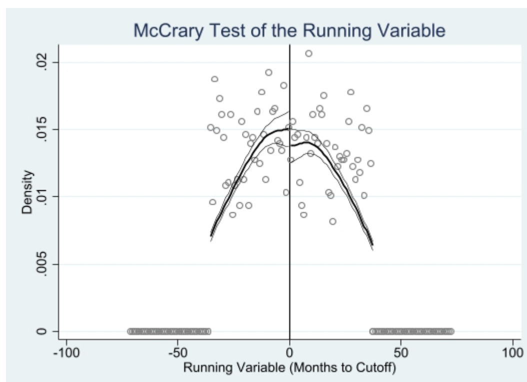
Figure 3.1: McCrary Test for the Sample Used in the First Outcome Variable



McCrary Test at a bandwidth of 6 months



McCrary Test at a bandwidth of 18 months



McCrary Test at a bandwidth of 36 months

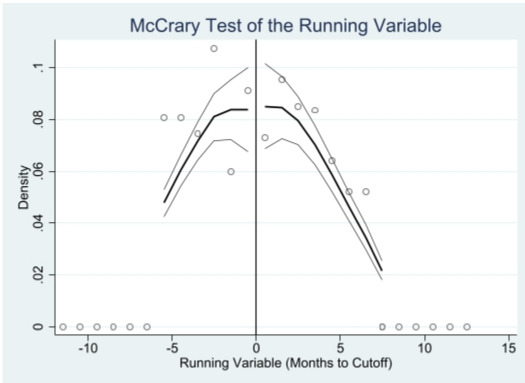
Table 3.3: McCrary test

Bandwidth	Discontinuity Estimate
6	0.028 (0.165)
18	-0.105 (0.097)
36	-0.089 (0.067)

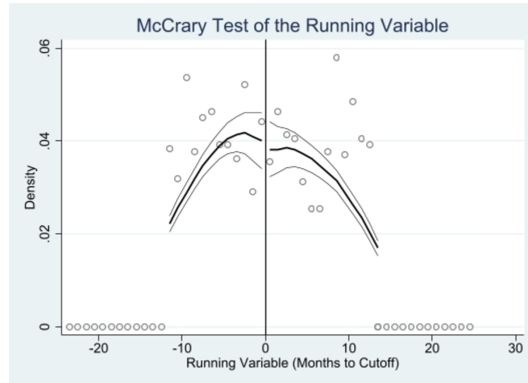
Standard Errors in parenthesis

36 Note that the cutoff of the McCrary test refers to the cohort with birth year- 2007 and birth month 12 instead of birthyear-2007 and birth month 11. This is because the study omits individuals with the birth year of 2007 and birth month of 2007. This adjustment is made because a threshold of that point would lead to zero observations at the cutoff, naturally giving rise to an upward/ downward discontinuity, and is not of interest.

Figure 3.2: McCrary Test for the Sample Used in the Second Outcome Variable



McCrary Test at a bandwidth of 6 months



McCrary Test at a bandwidth of 12 months

Table 3.4: McCrary Test

Bandwidth	Discontinuity Estimate
6	0.012 (0.168)
12	-0.031 (0.122)

Standard Errors in parenthesis

I then move forward towards determining the continuity of pre-determined covariates that might affect the outcome variables (for the combined sample of the first born and the second born girl). I do this by substituting the outcome variables by the pre-determined variables in the regression models. I test for the age of the mother, religious backgrounds, whether the household resides in an urban area, education of the mother, and household wealth.

Table 3.5 shows the continuity at different bandwidths for the first sample. I witness a weak discontinuity in religion of the household. However, the discontinuity is non-robust across the bandwidths. Appendix (C) section 1 shows the continuity of the pre-determined variables. The discontinuities again stand non-robust for the sample formed for education outcomes as shown in Table 3.6 and Appendix C section 2.

Table 3.5: Continuity of Covariates (Sample for Outcome 1: Whether the Child is Alive at the Time of Interview)

Variable	Bandwidth		
	6 months	18 months	36 months
Poor	-0.128	-0.025	0.002
	(0.081)	(0.040)	(0.026)
Middle Income	0.116	0.000	-0.019
	(0.074)	(0.036)	(0.023)
Rich	0.012	0.024	0.017
	(0.043)	(0.021)	(0.014)
Hindu	-0.091	-0.047	-0.022
	(0.075)	(0.039)	(0.027)
Muslim	0.091	0.040	0.017
	(0.075)	(0.039)	(0.027)
Urban	0.009	-0.041	-0.004
	(0.059)	(0.027)	(0.018)
Education of Woman	0.945	0.433	0.278
	(0.755)	(0.400)	(0.272)
Observations	679	2157	4110

Note: All estimates are weighted using state women's sampling weights; Robust standard errors in parentheses * p < 0.1, **p < 0.05, *** p < 0.01

Table 3.6: Continuity of Covariates (Sample for Outcome 2: Schooling Outcomes)

Variable	Bandwidth	
	6 months	12 months
Poor	-0.082	-0.057
	(0.083)	(0.052)
Middle Income	0.078	0.052
	(0.075)	(0.047)
Rich	0.003	0.005
	(0.045)	(0.027)
Hindu	-0.089	-0.098**
	(0.076)	(0.047)
Muslim	0.089	0.091*
	(0.076)	(0.047)
Urban	0.004	-0.034
	(0.061)	(0.036)
Education of Woman	0.793	0.217
	(0.771)	(0.508)
Observations	635	1325

Note: All estimates are weighted using state women's sampling weights; Robust standard errors in parentheses * p < 0.1, **p < 0.05, *** p < 0.01

3.4.2 Outcome Variables[37]

In this section, I discuss my regression estimates. I find no find no difference in the survival rate of the first-born girl child (Table 3.7) on either side of the cutoff. Using a 36-month bandwidth, with a 0.008 increase in survival for first-born girls, the upper bound of the 95% confidence interval is 0.043. Therefore, I can rule out effects larger than 0.043 with 95% confidence, which might make it harder to say that I find no meaningful effect. I find a non-robust significant increase in the survival rate of the second born child where the first born was a daughter (in the range of 5.3 to 6.4 percentage points) (Table 3.8).

I further study whether the policy led to an improvement in the education outcomes for the first-born girl. I find that the policy did lead to a consistent increase in the probability of ever attending school (Table 3.11) and the probability of attending school during the current school year (Table 3.10). However, the increase is insignificant. Similarly, eligible girls are likely to be in a higher grade in the current school year compared to the ineligible girls (Table 3.12). This might imply an earlier start at school. Furthermore, conditioning on the probability that the child ever went to school, I witness an increased number of schooling years of eligible children during the current school year (Table 3.13). There is no significant impact on schooling outcomes on the second born girl child.

Table 3.7: Probability the Child is Alive at the Time of the Interview (First Born Child)

	(1)	(2)	(3)
	Alive	Alive	Alive
Estimate	-0.044	0.003	0.008
	(0.035)	(0.023)	(0.018)
Mean	0.969	0.952	0.944
Bandwidth	6 months	18 months	36 months
Observations	363	1106	2141

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

[37] Sections 3 and 4 in the Appendix (C) show the graphical analysis of the outcome variables for the combined sample of the first born and second born girl.

Table 3.8: Probability the Child is Alive at the Time of the Interview (Second Born Girl When the First Born is a Female)

	(1) Alive	(2) Alive	(3) Alive
Estimate	0.092	0.064*	0.053**
	(0.074)	(0.033)	(0.024)
Mean	0.964	0.963	0.961
Bandwidth	6 months	18 months	36 months
Observations	156	551	1016

Note: All estimates are weighted using state women's sampling weights; Robust standard errors in parentheses * p < 0.1, **p < 0.05, *** p < 0.01

Table 3.9: Probability the Child is Alive at the Time of the Interview (Second Born Girl When the First Born is a Male)

	(1) Alive	(2) Alive	(3) Alive
Estimate	0.019	0.018	0.012
	(0.050)	(0.033)	(0.025)
Mean	0.971	0.970	0.952
Bandwidth	6 months	18 months	36 months
Observations	160	500	953

Note: All estimates are weighted using state women's sampling weights; Robust standard errors in parentheses * p < 0.1, **p < 0.05, *** p < 0.01

Table 3.10: Probability the Child Went to School During the Current School Year (First Born Child)

	(1) Attending School	(2) Attending School
Estimate	0.069	0.060
	(0.073)	(0.052)
Mean	0.905	0.881
Bandwidth	6 months	12 months
Observations	338	685

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table 3.11: Probability the Child Ever Went to School (First Born Child)

	(1) Child Ever Went to School	(2) Child Ever Went to School
Estimate	0.079	0.061
	(0.072)	(0.051)
Mean	0.909	0.884
Bandwidth	6 months	12 months
Observations	338	685

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table 3.12: Grade Attended During the Current School Year (First Born Child)

	(1) Grade Attended During the Current School Year	(2) Grade Attended During the Current School Year
Estimate	0.309	0.312
	(0.297)	(0.204)
Mean	1.747	1.815
Bandwidth	6 months	12 months
Observations	308	606

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table 3.13: Years of Education Conditional on Ever Attending School (First Born Child)

	(1) Years of Education Conditional on Ever Attending School	(2) Years of Education Conditional on Ever Attending School
Estimate	0.290	0.413*
	(0.297)	(0.217)
Mean	1.752	1.831
Bandwidth	6 months	12 months
Observations	310	610

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table 3.14: Probability the Child Went to School During the Current School Year (Second Born Girl When the First Born is a Female)

	(1) Attending School	(2) Attending School
Estimate	-0.000	0.031
	(0.105)	(0.084)
Mean	0.867	0.895
Bandwidth	6 months	12 months
Observations	141	296

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table 3.15: Probability the Child Ever Went to School (Second Born Girl When the First Born is a Female)

	(1) Child Ever Went to School	(2) Child Ever Went to School
Estimate	-0.005	0.027
	(0.104)	(0.084)
Mean	0.869	0.896
Bandwidth	6 months	12 months
Observations	141	296

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table 3.16: Grade Attended During the Current School Year (Second Born Girl When the First Born is a Female)

	(1) Grade Attended During the Current School Year	(2) Grade Attended During the Current School Year
Estimate	-0.644	0.219
	(0.442)	(0.305)
Mean	1.756	1.849
Bandwidth	6 months	12 months
Observations	123	265

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table 3.17: Years of Education Conditional on Ever Attending School (Second Born Girl When the First Born is a Female)

	(1) Years of Education Conditional on Ever Attending School	(2) Years of Education Conditional on Ever Attending School
Estimate	-0.635	0.228
	(0.442)	(0.304)
Mean	1.752	1.847
Bandwidth	6 months	12 months
Observations	124	266

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table 3.18: Probability the Child Went to School During the Current School Year (Second Born Girl When the First Born is a Male)

	(1) Attending School During Current School Year	(2) Attending School During Current School Year
Estimate	-0.011	-0.004
	(0.128)	(0.085)
Mean	0.881	0.085
Bandwidth	6 months	12 months
Observations	156	344

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table 3.19: Probability the Child Ever Went to School (Second Born Girl When the First Born is a Male)

	(1) Child Ever Went to School	(2) Child Ever Went to School
Estimate	-0.060	-0.015
	(0.122)	(0.080)
Mean	0.900	0.854
Bandwidth	6 months	12 months
Observations	156	344

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table 3.20: Grade Attended During the Current School Year (Second Born Girl When the First Born is a Male)

	(1) Grade Attended During the Current School Year	(2) Grade Attended During the Current School Year
Estimate	-0.130	0.067
	(0.568)	(0.299)
Mean	1.848	1.841
Bandwidth	6 months	12 months
Observations	139	294

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table 3.21: Years of Education Conditional on Ever Attending School (Second Born Girl When the First Born is a Male)

	(1) Years of Education Conditional on Ever Attending School	(2) Years of Education Conditional on Ever Attending School
Estimate	-0.032	0.109
	(0.564)	(0.299)
Mean	1.808	1.842
Bandwidth	6 months	12 months
Observations	142	297

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

3.4.3 Robustness Check

In this section, I discuss the robustness of results to a flexible linear (Appendix C Section 5), flexible quadratic (Appendix C Section 6) and local quadratic (Appendix C Section 7) RD specifications. I further add baseline household characteristics to the original specification (Appendix C Section 8). While my local linear and quadratic specifications include interaction terms between the dummy variable, D with polynomial terms of the running variable, the flexible specifications do not include an interaction term.

My estimates stand robust for the survival probability of the second born girl child where the first born is a female for most of the specifications. However, the estimates lose significance when I consider the number of schooling years considering only the children who ever went to school.

3.5 Conclusion

This study studies the impact of a policy aimed at curbing female feticide. I study the effect of the policy on the survival rate of the girl child (as measured by whether the child was alive at the time of the interview) and their potential education outcomes.

I find no significant effects of the policy on the survival rate of the first-born girl child. However, I find some suggestive evidence of a higher probability of survival for the second born girl child where the first-born child is a female. This result is robust across alternative specifications.

Furthermore, the results indicate that while such policies may help already enrolled children (first born) children stay longer in school, they do not affect the probability of school attendance per se. My results for schooling outcomes are consistent with other studies (Sinha and Yoong 2009) studying the impact of similar long-term financial incentives in India.

The results should be interpreted with caution since my study does not compute a first stage. This implies that the estimates computed in the study are downward biased.

The small effects from the policy could be due to the way the policy is formulated. Firstly, one of the problems could be the fact that the policy requires the family to hold a BPL card. NFHS divides the population into poorest, poorer, middle, richer, and the richest groups. I find that the probability of holding a BPL card for these groups are 70.56 percent; 67.95 percent; 62.95 percent, 49.73

percent, and 25.90 percent respectively. This supports some earlier research regarding some rich families holding the BPL card. Hence, although the probability of holding a BPL card reduces with wealth, there could still be a possibility that the family is not holding a BPL card or other required paperwork required for policy eligibility.

Secondly, a one-time long cash transfer after 18 years of childbirth might not be suitable for encouraging birth of a girl. Some periodic incentive might be helpful (Sekher 2010). Furthermore, Sekher (2010) also finds that there was a lack of adequate publicity for the policy and the state ran out of funds. This could imply that many people wanting to enroll could not enroll in the policy.

One of the merits of the policy has been celebrated in terms of increased birth registration. According to NFHS-4, there has been an increase in the birth registration from 5.8 (NFHS – 2005-2006) to 60.7 (NFHS 2015-2016).[38] My results might suggest that while the policy could have increased a low-cost activity like birth registration, it might not improve high-cost activities like giving birth to a girl child and investing in her per se.

[38] https://state.bihar.gov.in/

Bibliography

Amuedo-Dorantes, Catalina & Antman, Francisca. (2017): "Schooling and Labor Market Effects of Temporary Authorization: Evidence from DACA. Journal of Population Economics, 30.

Anand, S., & Gaur, R. (2013). Promoting the survival of girl children through conditional cash transfers: Is it a sustainable approach?. Indian Anthropologist, 55-71.

Andreyeva T, Long MW, Brownell KD. The Impact of Food Prices on Consumption: A Systematic Review of Research on the Price Elasticity of Demand for Food. American Journal of Public Health. 2010-02-01 2010;100(2):216-222.

Angrist J., and Krueger A. (1991):"Does Compulsory School Attendance Affect Schooling and Earnings?" The Quarterly Journal of Economics, 106(4), 979–1014.

Anukriti, S. 2018. "Financial Incentives and the Fertility-Sex Ratio Trade-Off." American Economic Journal: Applied Economics, 10 (2): 27-57.

Bae Jung, 2020. "Immigration Relief and Insurance Coverage: Evidence from Deferred Action for Childhood Arrivals," The B.E. Journal of Economic Analysis & Policy, De Gruyter, vol. 20(3), pages 1-37, July.

Battistin, Erich, Agar Brugiavini, Enrico Rettore, and Guglielmo Weber. 2009. "The Retirement Consumption Puzzle: Evidence from a Regression Discontinuity Approach." American Economic Review, 99 (5): 2209-26

Bennett R, Zorbas C, Huse O, et al. Prevalence of healthy and unhealthy food and beverage price promotions and their potential influence on shopper purchasing behaviour: A systematic review of the literature. Obesity Reviews. 2020;21(1):e12948.

Buckles K., and Hungerman D. (2013): "Season of Birth and Later Outcomes: Old Questions, New Answers.," The review of economics and statistics, 95, 711-724.

Catalina Amuedo-Dorantes, Francisca Antman, Can authorization reduce poverty among undocumented immigrants? Evidence from the Deferred Action for Childhood Arrivals program, Economics Letters,Volume 147, 2016, Pages 1-4, ISSN 0165-1765, https://doi.org/10.1016/j.econlet.2016.08.001.

Emily Battaglia, Did DACA Harm US-Born Workers? Temporary Work Visas and Labor Market Competition, Journal of Urban Economics, Volume 134, 2023, 103512, ISSN 0094-1190, https://doi.org/10.1016/j.jue.2022.103512.

Feeding America, . Annual Reports and Financials; 2022.

García-Pérez, M. DACA Recipients and Their Health Insurance Dream: Employment, Schooling, and Health Coverage. J Econ Race Policy 2, 77–108 (2019). https://doi.org/10.1007/s41996-019-00029-x

Gelman, A., and G. Imbens. 2016. "Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs." Unpublished manuscript.

Giuntella, Osea, and Jakub Lonsky. "The effects of DACA on health insurance, access to care, and health outcomes." Journal of Health Economics 72 (2020): 102320.

Grembi, Veronica, Tommaso Nannicini, and Ugo Troiano. 2016. "Do Fiscal Rules Matter?" American Economic Journal: Applied Economics, 8 (3): 1-30.

Gucciardi E, Vahabi M, Norris N, Del Monte JP, Farnum C. The Intersection between Food Insecurity and Diabetes: A Review. Current Nutrition Reports. 2014-12-01 2014;3(4):324-332.

Haghpour B, Sahabeh E, Halvari H. Opportunity cost in consumer behavior: Definitions, operationalizations, and ambiguities. International Journal of Consumer Studies. 2022-09-01 2022;46(5):1942-1959.

Hamilton, E. R., Patler, C., and Savinar, R. (2021). Transition into liminal legality: Dacas mixed impacts on education and employment among young adult immigrants in california. Social Problems, 68(3):675–695.

Healthy Eating Research. Healthy Eating Research Nutrition Guidelines for the Charitable Food System 2020.

Hoefer, M., Rytina, N. and Baker, B. (2012): "Estimates of the Unauthorized Immigrant Population Residing in the United States: January 2011." Population Estimates.

Hsin A., and Ortega F. (2018): "The Effects of Deferred Action for Childhood Arrivals on the Educational Outcomes of Undocumented Students," Demography, Springer; Population Association of America (PAA), 55(4), 1487-1506.

Jain, M. (2022). Do Financial Incentives Affect Fertility and Sex Ratio at Birth? An Impact Evaluation of Madhya Pradesh's Ladli Laxmi Yojana, working paper, University of Wisconsin-Madison.

Jayachandran, Seema, and Rohini Pande. 2017. "Why Are Indian Children So Short? The Role of Birth Order and Son Preference." American Economic Review, 107 (9): 2600-2629.

Karpyn A, Mccallops K, Wolgast H, Glanz K. Improving Consumption and Purchases of Healthier Foods in Retail Environments: A Systematic Review. International Journal of Environmental Research and Public Health. 2020-10-16 2020;17(20):7524.

Kaushal, Neeraj. 2006. "Amnesty Programs and the Labor Market Outcomes of Undocumented Workers." Journal of Human Resources 41 (3). https://doi.org/10.3368/jhr.XLI.3.631.

Kuka, E., Na'ama S., and Kevin S. (2020): "Do Human Capital Decisions Respond to the Returns to Education? Evidence from DACA." American Economic Journal: Economic Policy, 12 (1), 293-324.

Kumar,Sneha & Sinha,Nistha, 2018. "Preventing More"Missing Girls": A Review of Policies to Tackle Son Preference," Policy Research Working Paper Series 8635, The World Bank.

Martin KS, Wolff M, Callahan K, Schwartz MB. Supporting Wellness at Pantries: Development of a Nutrition Stoplight System for Food Banks and Food Pantries. Journal of the Academy of Nutrition and Dietetics. 2018;119(4):553-559.

McCrary, J. (2008): "Manipulation of the running variable in the regression discontinuity design: A density test." Journal of Econometrics, Elsevier, 142(2), 698-714.

Mckee SL, Gurganus EA, Atoloye AT, Xu R, Martin K, Schwartz MB. Pilot testing an intervention to educate and promote nutritious choices at food pantries. Journal of Public Health. 2023-04-01 2023;31(4):521-528.

Münscher R, Vetter M, Scheuerle T. A Review and Taxonomy of Choice Architecture Techniques. Journal of Behavioral Decision Making. 2016;29(5):511-524.

Patler, Caitlin et al. (2019): "Uncertainty About DACA May Undermine Its Positive Impact On Health For Recipients And Their Children." Health affairs ,38(5) : 738-745

Pope, N.G. (2016):"The effects of dacamentation: the impact of deferred action for childhood arrivals on unauthorized immigrants" .J. Publ. Econ. 143, 98-114.

Rivera-Batiz, F. Undocumented workers in the labor market: An analysis of the earnings of legal and illegal Mexican immigrants in the United States. J Popul Econ 12, 91–116 (1999). https://doi.org/10.1007/s001480050092

Ross M, Campbell EC, Webb KL. Recent trends in the nutritional quality of food banks' food and beverage inventory: Case studies of six California food banks. Journal of hunger & environmental nutrition. 2013;8(3):294-309.

Second Harvest Kitchen. Client Choice and Shopping Models. Available at: https://secondharvestkitchen.org/client-choice-shopping-models/. Accessed July 10, 2023, 2023.

Sekher, T.V (2010). Special Financial Incentive Schemes for the Girl Child in India: A Review of Select schemes, United Nations Population Fund, New Delhi.

Sekher, T.V. and F. Ram (2015). Conditional Cash Transfers for Girls in India: Assessment of a Girl Child Promotion Scheme from Beneficiary Perspective (Report of a project study sponsored by UNFPA- India), International Institute for Population Sciences (IIPS), Mumbai.

Shaw SC, Ntani G, Baird J, Vogel CA. A systematic review of the influences of food store product placement on dietary-related outcomes. Nutrition reviews. 2020;78(12):1030-1045.

Sinha, Nistha & Yoong, Joanne, 2009. "Long-term financial incentives and investment in daughters: evidence from conditional cash transfers in north India," Policy Research Working Paper Series 4860, The World Bank.

Slapo H, Schjøll A, Karevold KI, Sandaker I, Lekhal S. Using choice architecture interventions to nudge costumers towards healthier food purchase in real-life grocery stores. A review. Surgery for Obesity and Related Diseases. 2019;15(10):S251-S252.

Steven Ruggles, Sarah Flood, Ronald Goeken, Megan Schouweiler and Matthew Sobek. IPUMS USA: Version 12.0 [dataset]. Minneapolis, MN: IPUMS, 2022. https://doi.org/10.18128/D010.V12.0

Tran, N. (2023). The effects of deferred action for childhood arrivals on labor market outcomes.

U.S. Department of Health and Human Services, U.S. Department of Agriculture. 2015-2020 Dietary Guidelines for Americans 2015.

Venkataramani, Atheendar S., Sachin J. Shah, Rourke O'Brien, Ichiro Kawachi, and Alexander C. Tsai. 2017. "Health consequences of the US Deferred Action for Childhood Arrivals (DACA) immigration programme: a quasi-experimental study." Lancet Public Health 2. https://doi.org/10.1016/S2468-2667(17)30047-6.

Waterlander WE, De Mul A, Schuit AJ, Seidell JC, Steenhuis IH. Perceptions on the use of pricing strategies to stimulate healthy eating among residents of deprived neighbourhoods: a focus group study. International Journal of Behavioral Nutrition and Physical Activity. 2010-01-01 2010;7(1):44.

Appendix A

Chapter 1

Table A1: Descriptive Statistics (Bandwidth: 16 Quarters)

Variable	Mean		Difference	t-statistic
	Treatment Group	Control Group		
Treatment Group				
Total Income (in $1000)	25,650.72	27,937.73	-2,287.01	-5.10
Work	0.81	0.79	0.02	3.96
Hours of Work	32.42	31.71	0.71	2.43
School Attendance	0.05	0.05	0.00	0.78
Male	0.54	0.52	0.02	3.01
Age of Entry in the US	8.81	19.02	-10.21	-117.11
Cognitive Difficulty	0.02	0.03	-0.01	-5.30
Ambulatory Difficulty	0.01	0.02	-0.01	-4.68
Mobility Difficulty	0.01	0.02	-0.01	-4.56
Self-Care Difficulty	0.01	0.01	-0.00	-1.99
Hearing/Seeing Difficulty	0.02	0.02	-0.00	-1.48
Observations	5,731	124,718		

Table A2: Descriptive Statistics (Bandwidth: 24 Quarters)

Variable	Mean		Difference	t-statistic
	Treatment Group	Control Group		
Treatment Group				
Total Income (in $1000)	25,795.66	27,887.08	-2,091.42	-5.75
Work	0.82	0.80	0.02	4.76
Hours of Work	32.42	31.78	0.64	2.74
School Attendance	0.07	0.06	0.01	1.81
Male	0.54	0.52	0.02	3.71
Age of Entry in the US	8.76	19.05	-10.29	-147.28
Cognitive Difficulty	0.02	0.03	-0.01	-7.95
Ambulatory Difficulty	0.01	0.02	-0.01	-5.33
Mobility Difficulty	0.01	0.02	-0.01	-5.37
Self-Care Difficulty	0.01	0.01	-0.00	-2.94
Hearing/Seeing Difficulty	0.02	0.02	-0.00	-0.76
Observations	9,224	186,404		

Table A3: Estimates for Continuity of Covariates for Disability Measures (for the Baseline RD)

Outcome Variable:	Cognitive Difficulty			Ambulatory Difficulty			Mobility Difficulty			Self-Care Difficulty			Hearing./Seeing Difficulty		
Estimate	-0.014*	-0.010	-0.011	0.007	0.005	0.005	-0.002	-0.000	-0.001	-0.000	0.001	0.001	-0.001	0.001	-0.001
	(0.009)	(0.008)	(0.007)	(0.007)	(0.007)	(0.006)	(0.007)	(0.007)	(0.006)	(0.006)	(0.005)	(0.005)	(0.009)	(0.008)	(0.008)
Bandwidth	16	20	24	16	20	24	16	20	24	16	20	24	16	20	24

Note: All estimates are weighted using sampling weights; Robust standard errors in parentheses The analysis is performed using (2) . * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

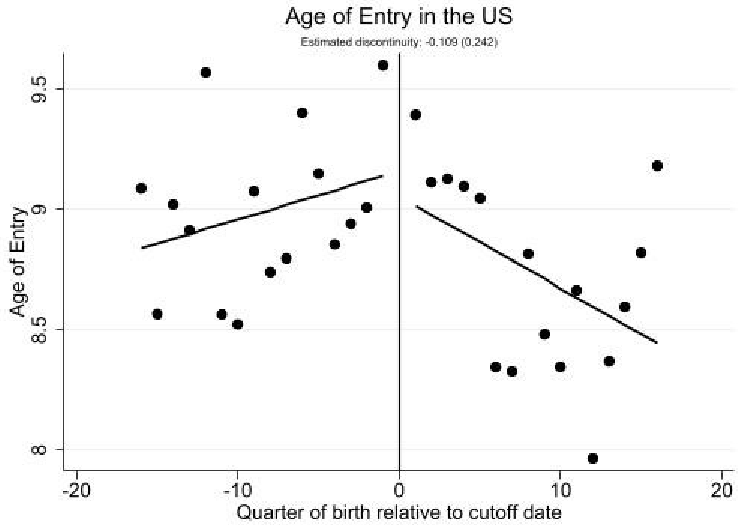


Figure A1: Covariate Smoothness Test of “Age of Entry” in the treatment group in the post DACA period (2014-2017) at a bandwidth of 16 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

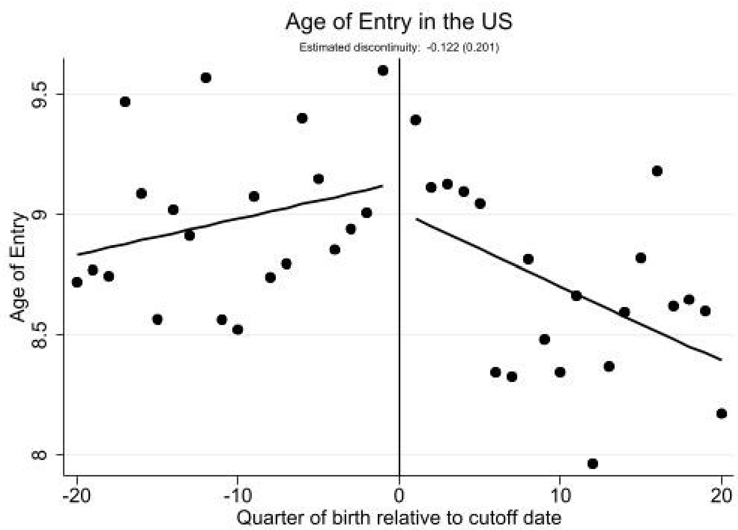


Figure A2: Covariate Smoothness Test of “Age of Entry” in the treatment group in the post DACA period (2014-2017) at a bandwidth of 20 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

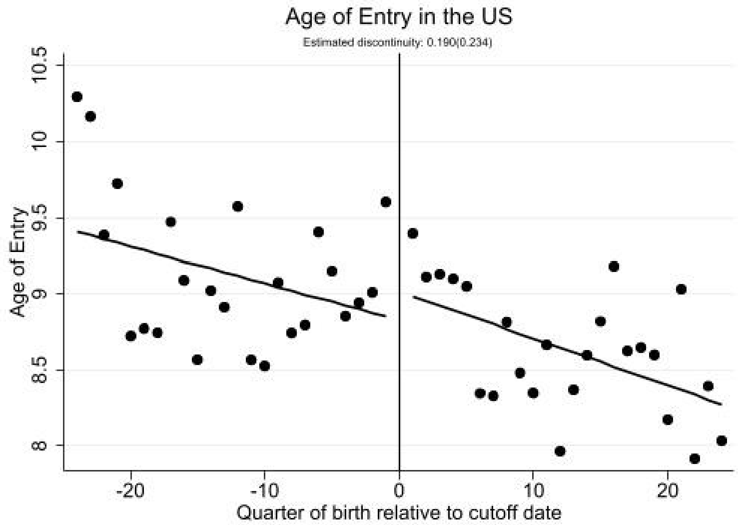


Figure A3: Covariate Smoothness Test of "Age of Entry" in the treatment group in the post DACA period (2014-2017) at a bandwidth of 24 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

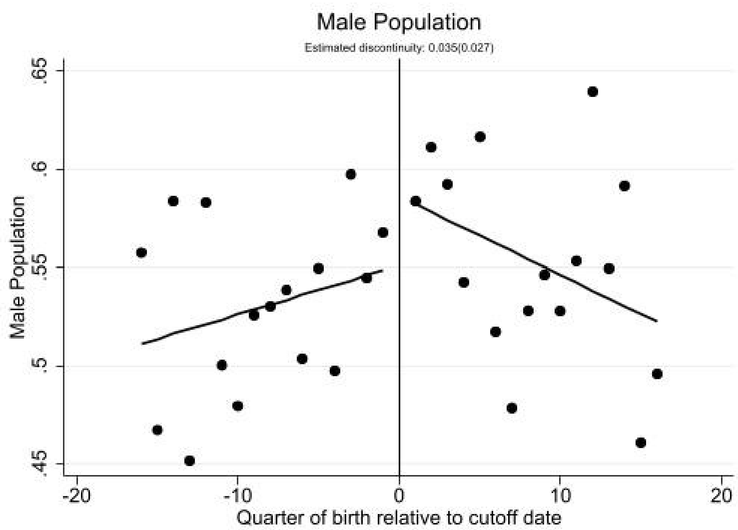


Figure A4: Covariate Smoothness Test of "Gender" in the treatment group in the post DACA period (2014-2017) at a bandwidth of 16 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

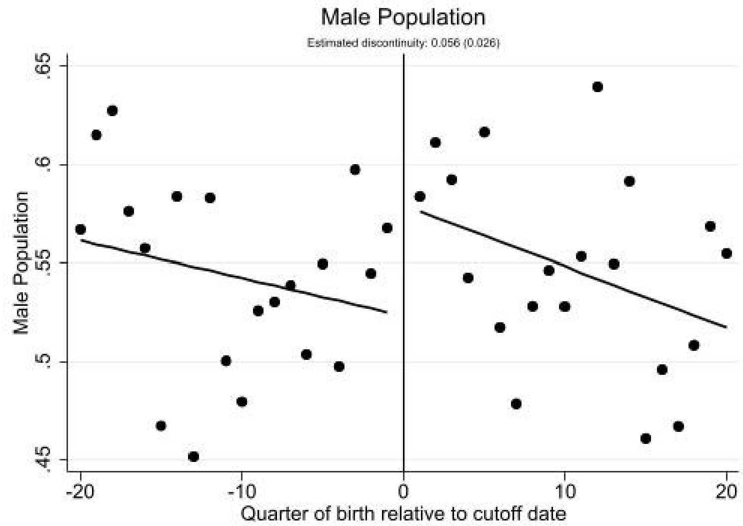


Figure A5: Covariate Smoothness Test of “Gender” in the treatment group in the post DACA period (2014-2017) at a bandwidth of 20 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

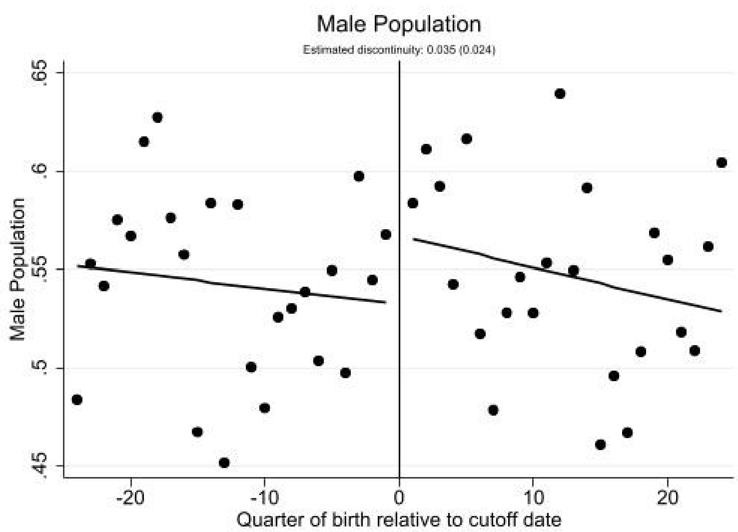


Figure A6: Covariate Smoothness Test of “Gender” in the treatment group in the post DACA period (2014-2017) at a bandwidth of 24 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

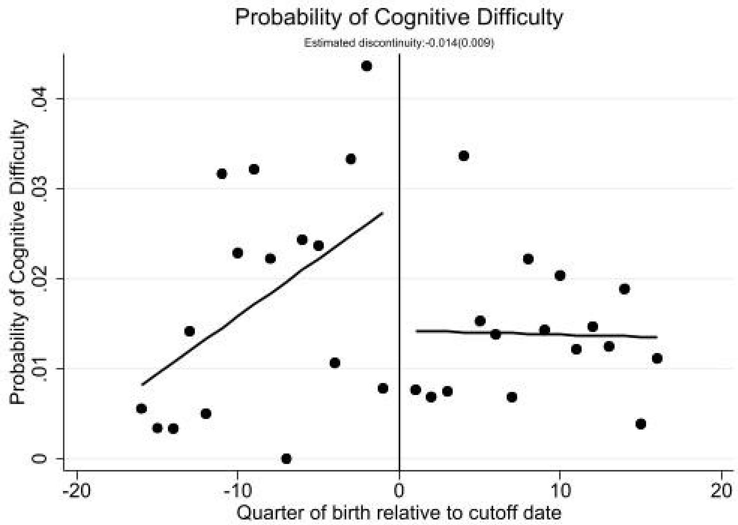


Figure A7: Covariate Smoothness Test of “Cognitive Difficulty” in the treatment group in the post DACA period (2014-2017) at a bandwidth of 16 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

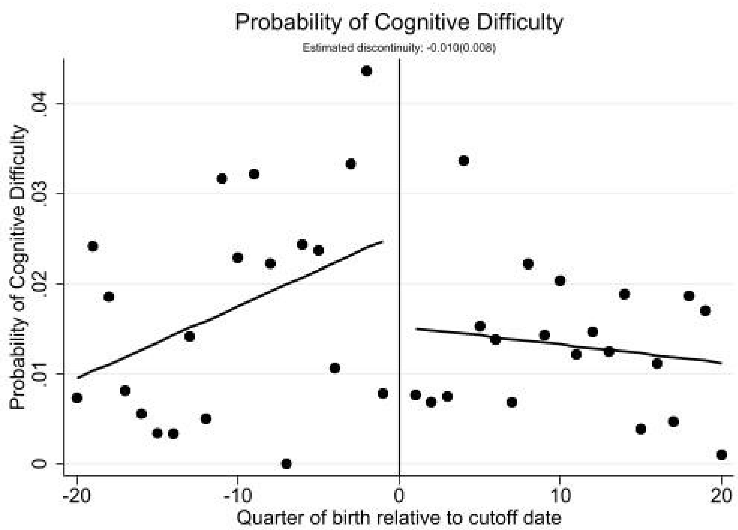


Figure A8: Covariate Smoothness Test of “Cognitive Difficulty” in the treatment group in the post DACA period (2014-2017) at a bandwidth of 20 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

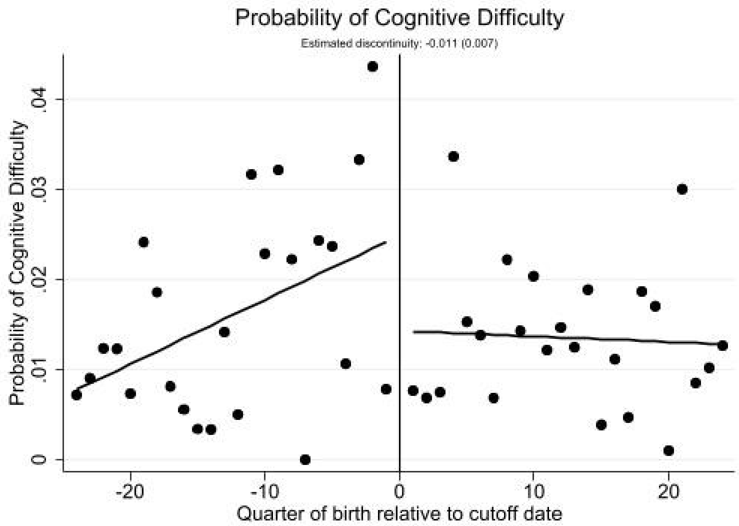


Figure A9: Covariate Smoothness Test of “Cognitive Difficulty” in the treatment group in the post DACA period (2014-2017) at a bandwidth of 24 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

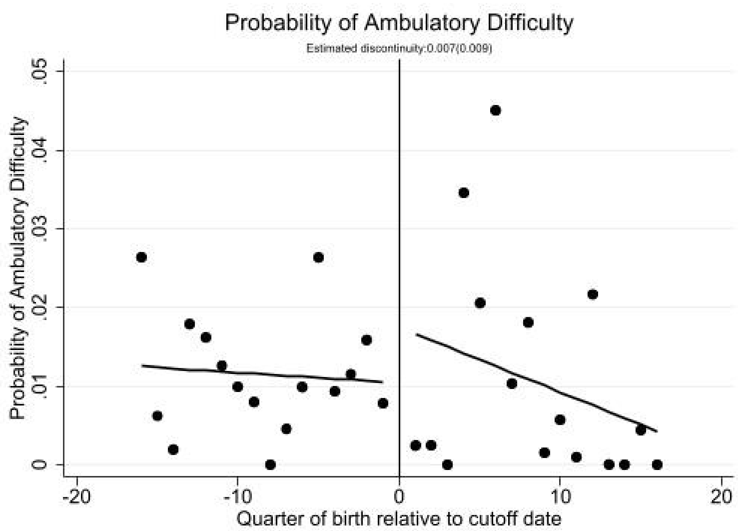


Figure A10: Covariate Smoothness Test of “Ambulatory Difficulty” in the treatment group in the post DACA period (2014-2017) at a bandwidth of 16 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

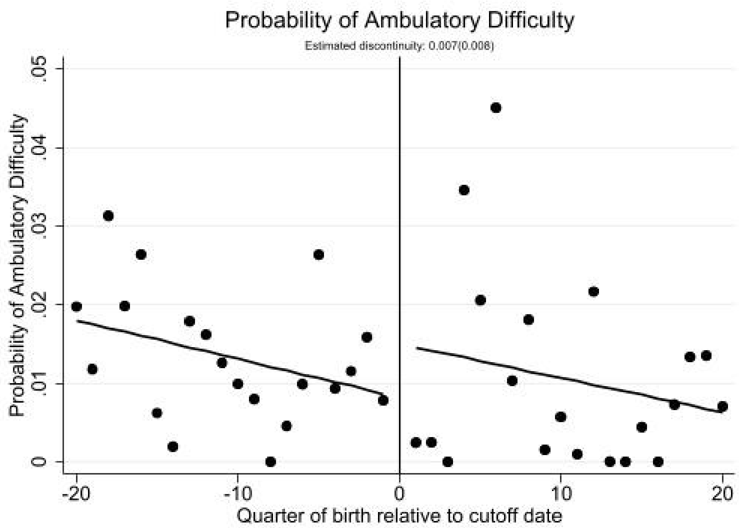


Figure A11: Covariate Smoothness Test of "Ambulatory Difficulty" in the treatment group in the post DACA period (2014-2017) at a bandwidth of 20 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

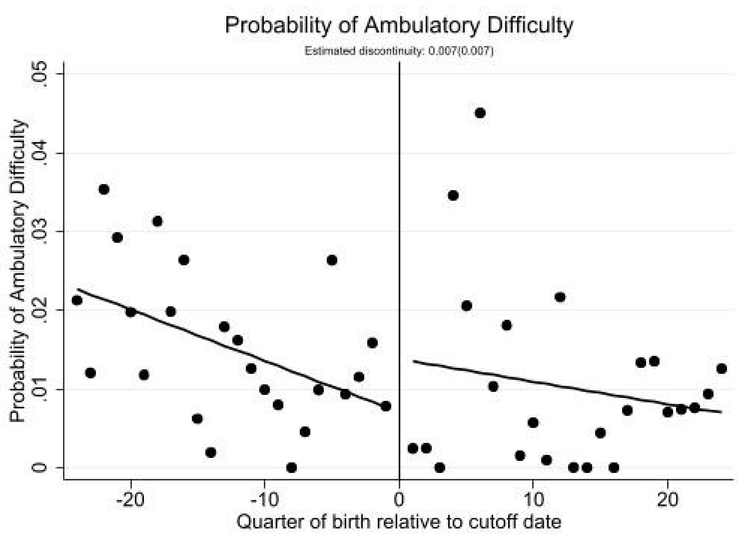


Figure A12: Covariate Smoothness Test of "Ambulatory Difficulty" in the treatment group in the post DACA period (2014-2017) at a bandwidth of 24 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

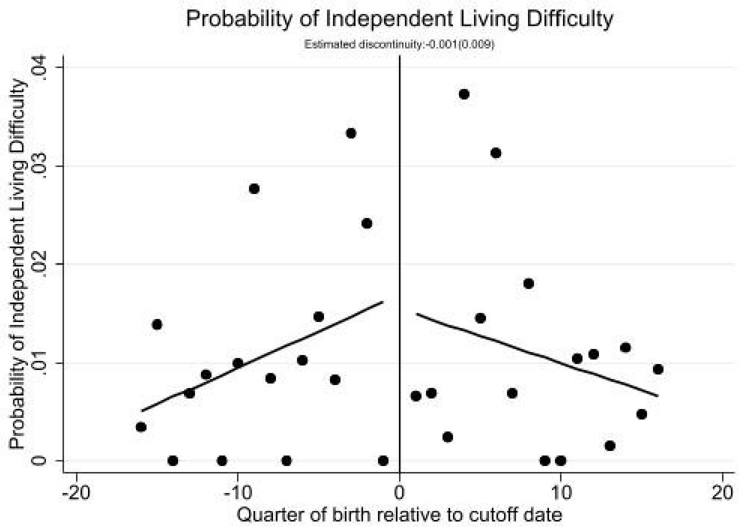


Figure A13: Covariate Smoothness Test of "Independent Living Difficulty" in the treatment group in the post DACA period (2014- 2017) at a bandwidth of 16 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

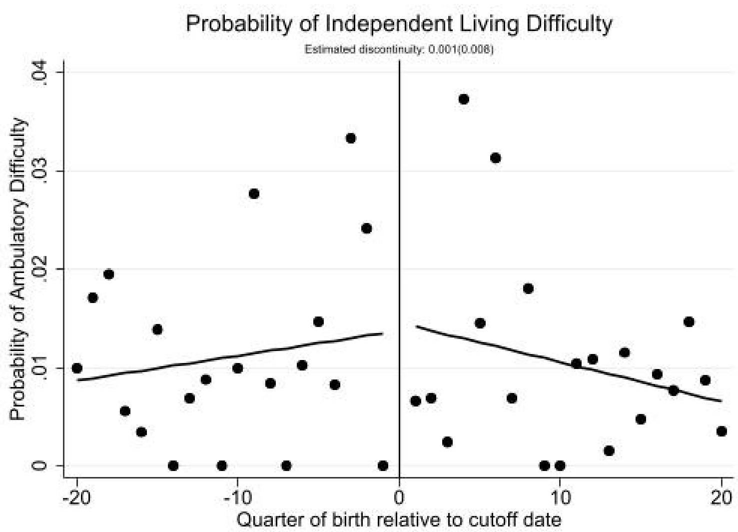


Figure A14: Covariate Smoothness Test of "Independent Living Difficulty" in the treatment group in the post DACA period (2014- 2017) at a bandwidth of 20 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

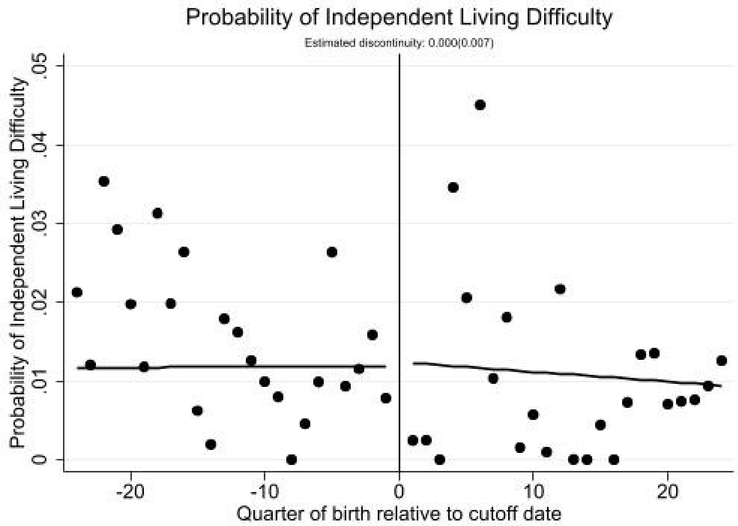


Figure A15: Covariate Smoothness Test of "Independent Living Difficulty" in the treatment group in the post DACA period (2014- 2017) at a bandwidth of 24 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

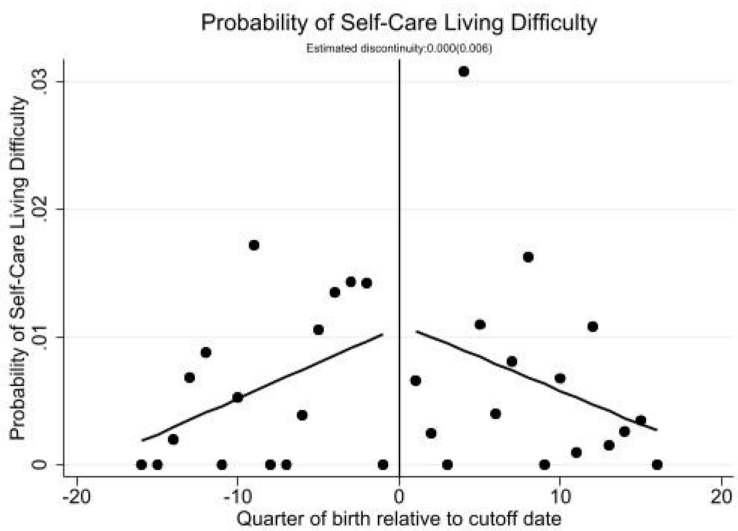


Figure A16: Covariate Smoothness Test of "Self-Care Difficulty" in the treatment group in the post DACA period (2014-2017) at a bandwidth of 16 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

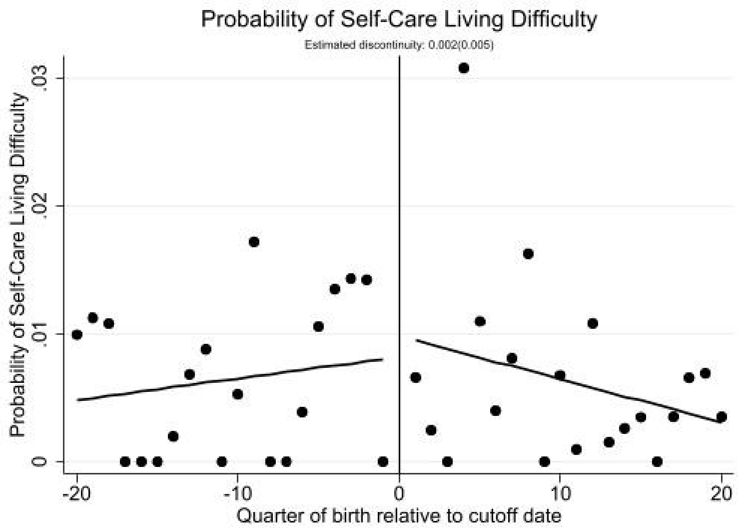


Figure A17: Covariate Smoothness Test of "Self-Care Difficulty" in the treatment group in the post DACA period (2014-2017) at a bandwidth of 20 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

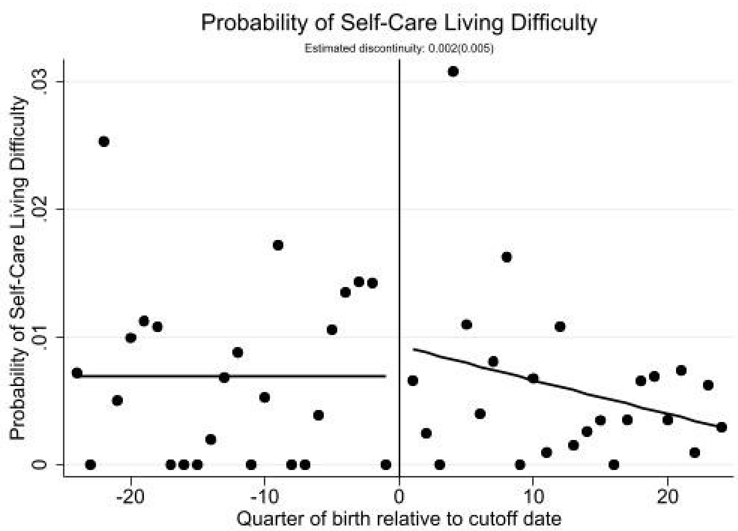


Figure A18: Covariate Smoothness Test of "Self-Care Difficulty" in the treatment group in the post DACA period (2014-2017) at a bandwidth of 24 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

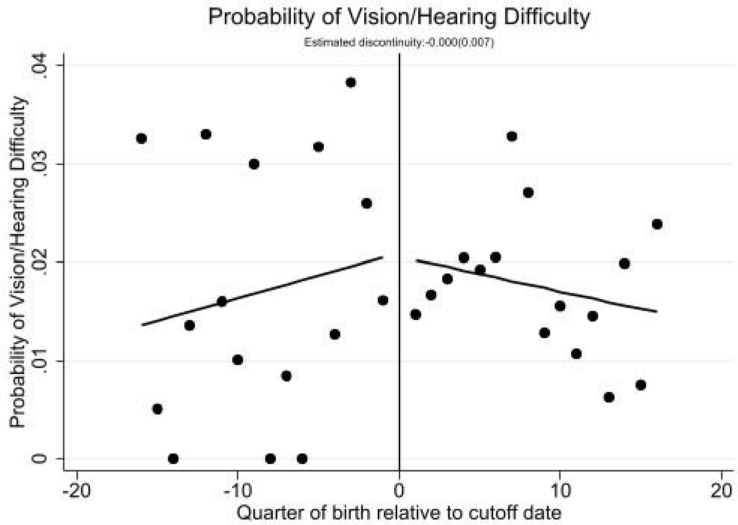


Figure A19: Covariate Smoothness Test of "Vision/Hearing Difficulty" in the treatment group in the post DACA period (2014-2017) at a bandwidth of 16 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

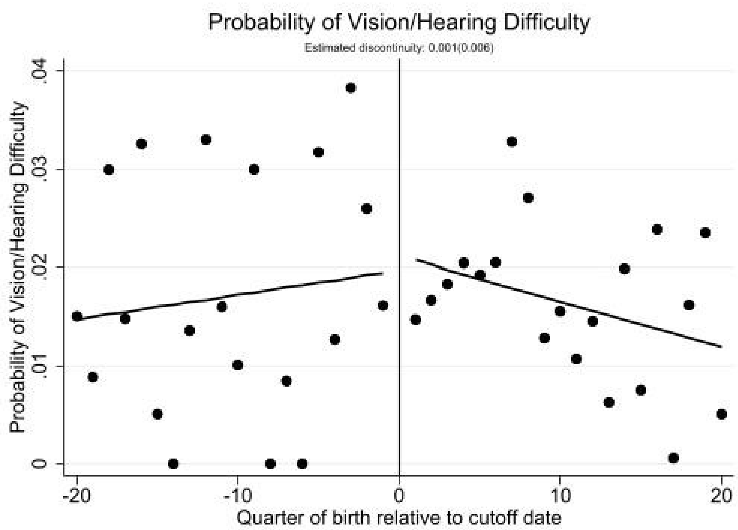


Figure A20: Covariate Smoothness Test of "Vision/Hearing Difficulty" in the treatment group in the post DACA period (2014-2017) at a bandwidth of 20 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

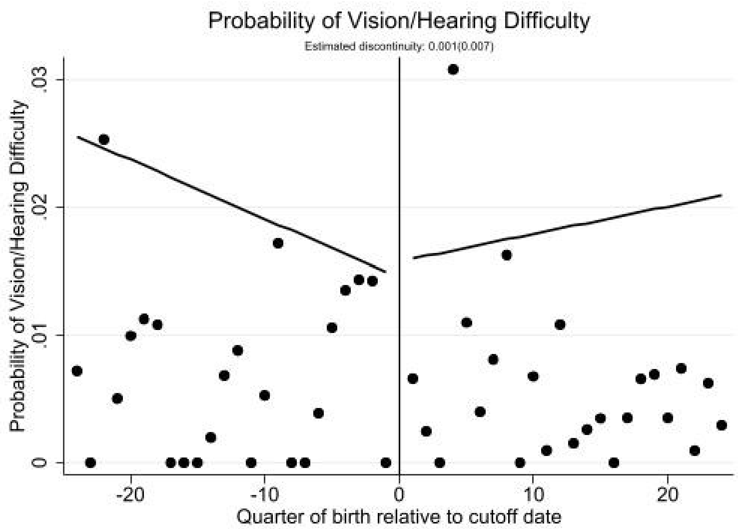


Figure A21: Covariate Smoothness Test of “Vision/Hearing Difficulty” in the treatment group in the post DACA period (2014-2017) at a bandwidth of 24 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

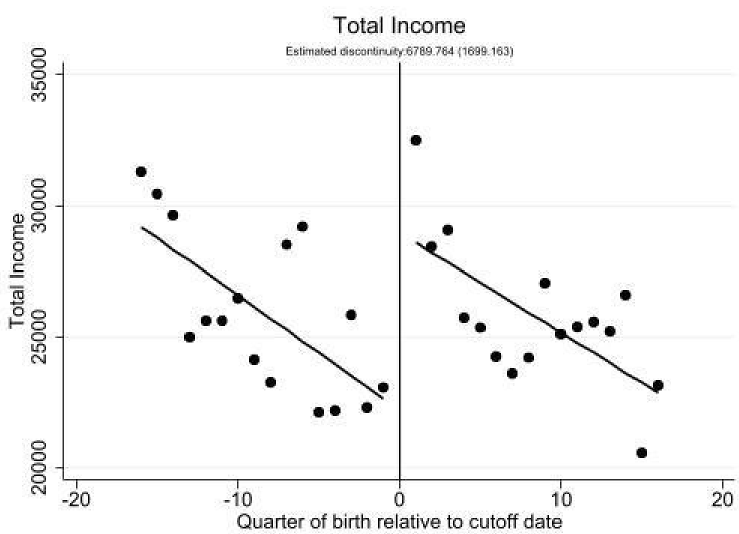


Figure A22: Total Income of the treatment group in the post DACA period (2014-2017) at a bandwidth of 16 quarters. The estimated discontinuity is computed using linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

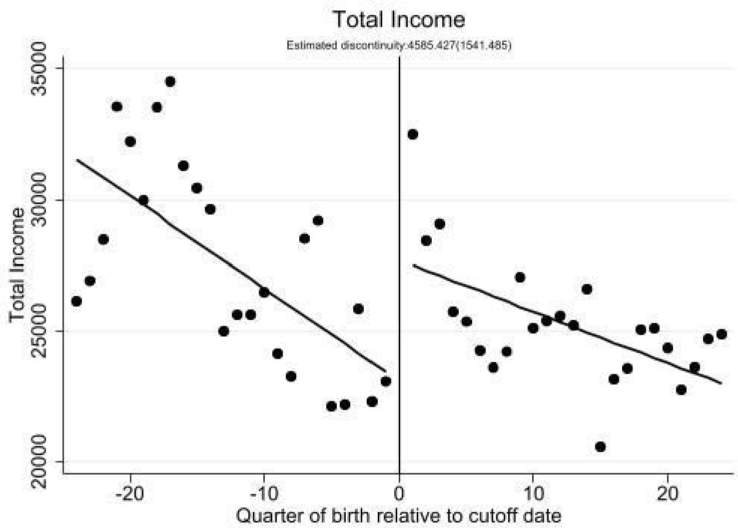


Figure A23: Total Income of the treatment group in the post DACA period (2014-2017) at a bandwidth of 24 quarters. The estimated discontinuity is computed using linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

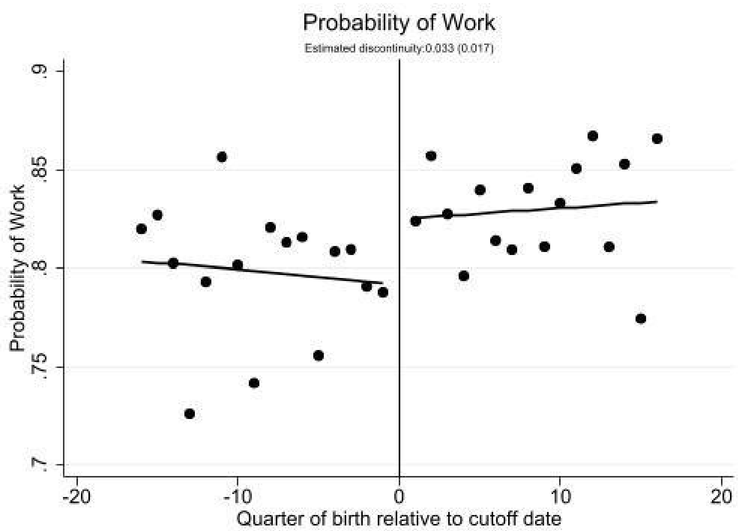


Figure A24: Probability of Work of the treatment group in the post DACA period (2014-2017) at a bandwidth of 16 quarters. The estimated discontinuity is computed using linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

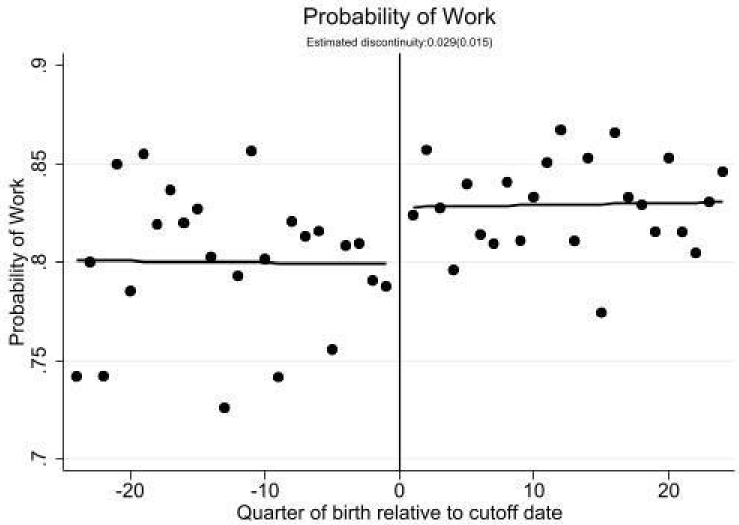


Figure A25: Probability of Work of the treatment group in the post DACA period (2014-2017) at a bandwidth of 24 quarters. The estimated discontinuity is computed using linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

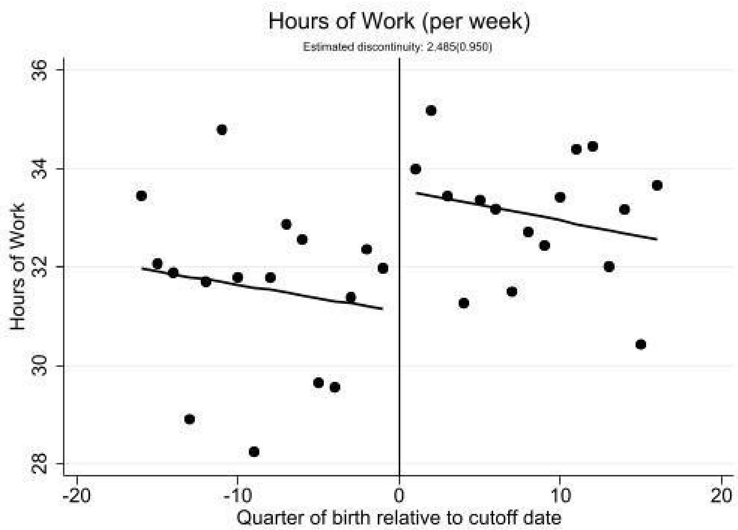


Figure A26: Hours of Work of the treatment group in the post DACA period (2014-2017) at a bandwidth of 16 quarters. The estimated discontinuity is computed using linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

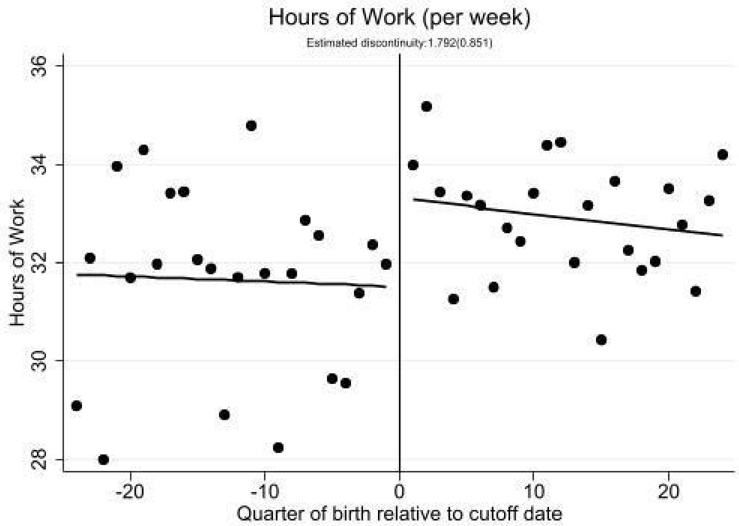


Figure A27: Hours of Work of the treatment group in the post DACA period (2014-2017) at a bandwidth of 24 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

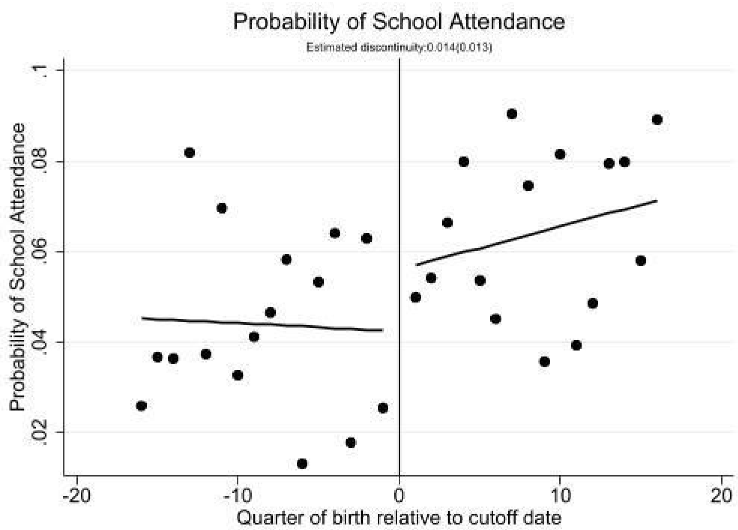


Figure 28: Probability of School Attendance of the treatment group in the post DACA period (2014-2017) at a bandwidth of 16 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

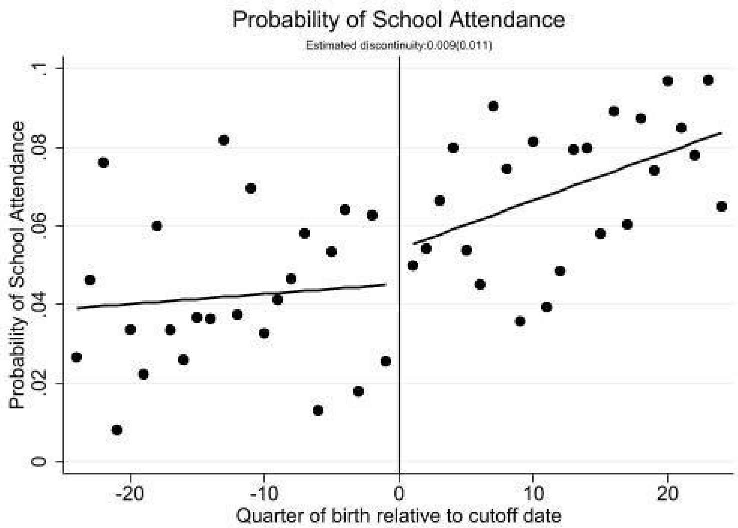


Figure A29: Probability of School Attendance of the treatment group in the post DACA period (2014-2017) at a bandwidth of 24 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

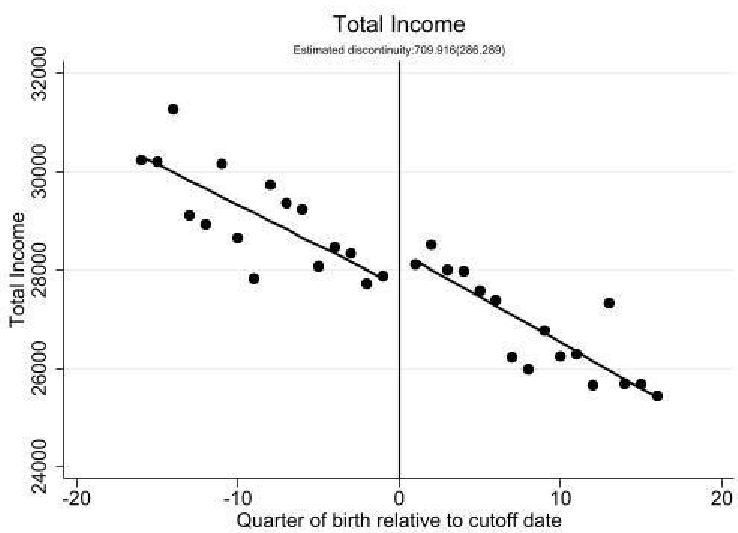


Figure A30: Total Income of the control group in the post DACA period (2014-2017) at a bandwidth of 16 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

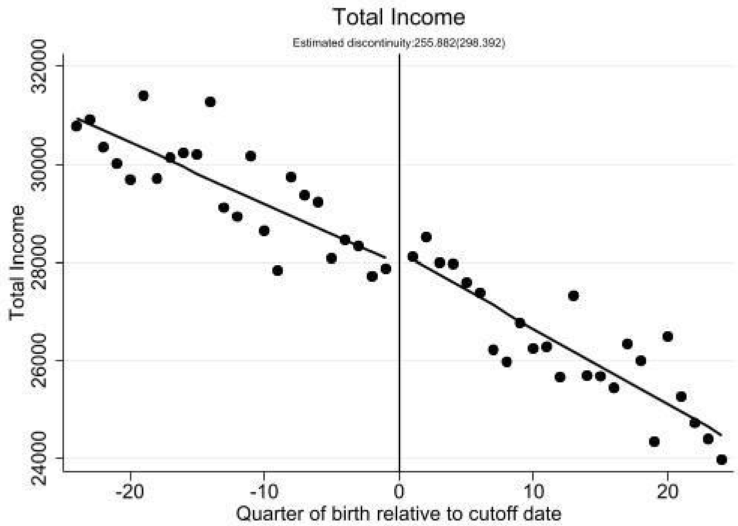


Figure A31: Total Income of the control group in the post DACA period (2014-2017) at a bandwidth of 24 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

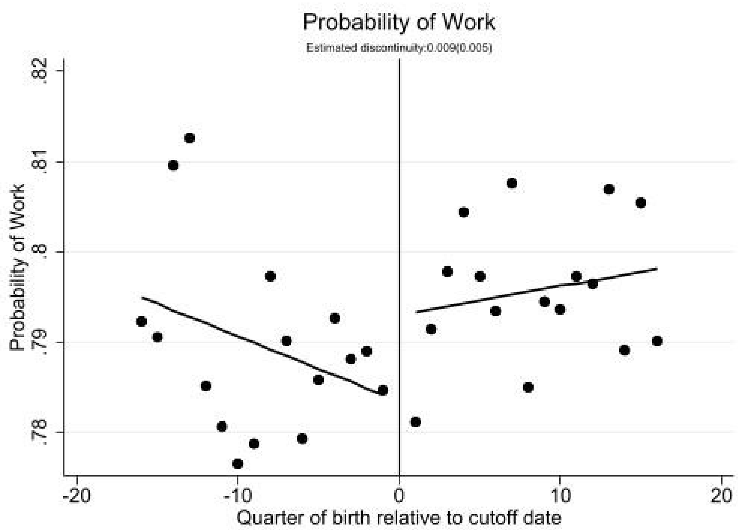


Figure A32: Probability of Work of the control group in the post DACA period (2014-2017) at a bandwidth of 16 quarters. The estimated discontinuity is computed using linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

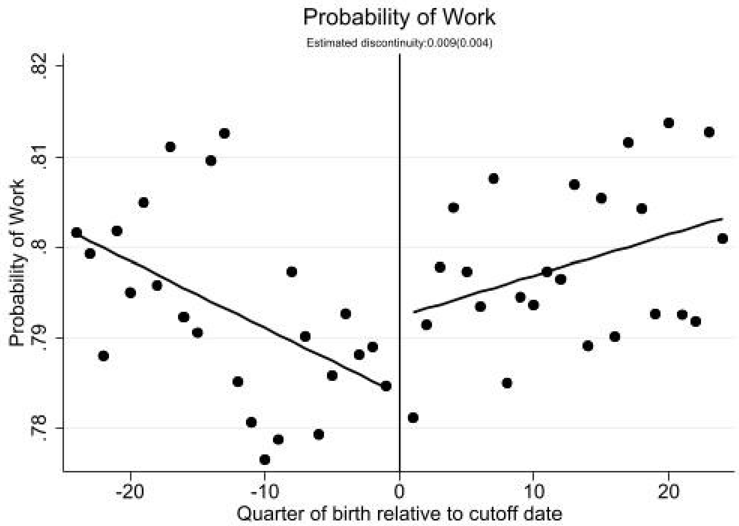


Figure A33: Probability of Work of the control group in the post DACA period (2014-2017) at a bandwidth of 24 quarters. The estimated discontinuity is computed using linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

Figure A34: Hours of Work of the control group in the post DACA period (2014-2017) at a bandwidth of 16 quarters. The estimated discontinuity is computed using linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

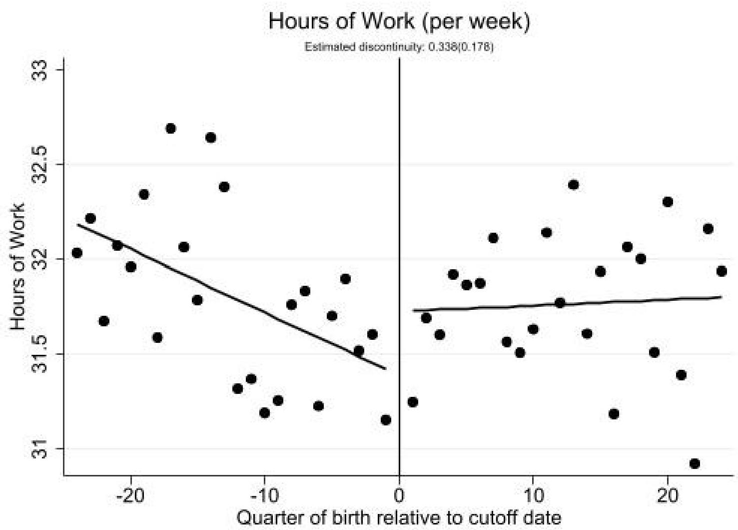


Figure A35: Hours of Work of the control group in the post DACA period (2014-2017) at a bandwidth of 24 quarters. The estimated discontinuity is computed using linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

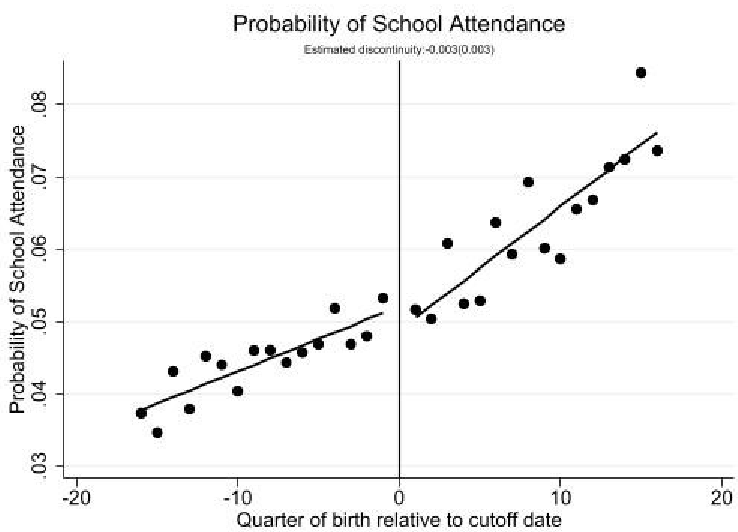


Figure A36: Probability of School Attendance of the control group in the post DACA period (2014-2017) at a bandwidth of 16 quarters. The estimated discontinuity is computed using linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

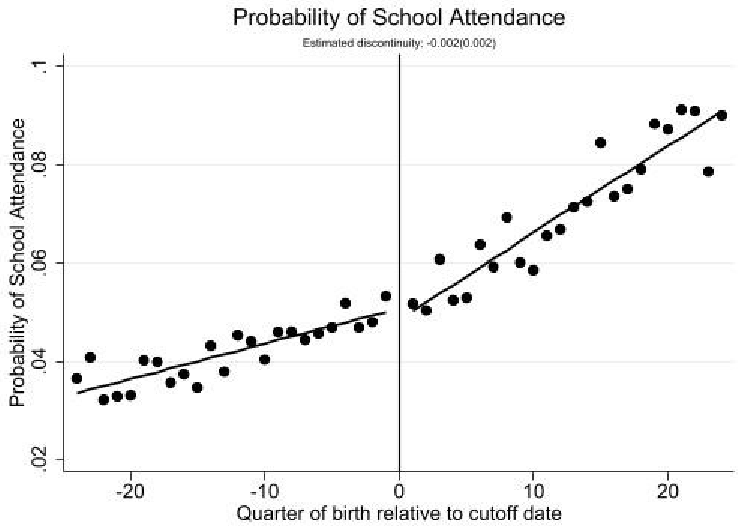


Figure A37: Probability of School Attendance of the control group in the post DACA period (2014-2017) at a bandwidth of 24 quarters. The estimated discontinuity is computed using linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

Table A4: Descriptive Statistics; Sample: Entire Population (Bandwidth: 16 Quarters)

Variable	Mean		Difference	t-statistic
	Treatment Group	Control Group		
Treatment Group				
Total Income (in $1000)	25650.63	41977.68	-16327.05	-37.18
Work	0.82	0.83	-0.01	-2.15
Hours of Work	32.42	33.67	-1.25	-4.40
School Attendance	0.05	0.07	-0.02	-4.97
Male	0.54	0.50	0.04	5.39
Age of Entry in the US	8.81	19.98	-11.17	-135.15
Cognitive Difficulty	0.02	0.04	-0.02	-12.68
Ambulatory Difficulty	0.01	0.02	-0.01	-8.37
Mobility Difficulty	0.01	0.03	-0.02	-10.08
Self-Care Difficulty	0.01	0.01	-0.00	-3.82
Hearing/Seeing Difficulty	0.02	0.02	-0.00	-2.38
Observations	5,731	1,187,288		

Table A5: Descriptive Statistics; Sample: Entire Population (Bandwidth: 20 Quarters)

Variable	Mean		Difference	t-statistic
	Treatment Group	Control Group		
Treatment Group				
Total Income (in $1000)	25925.02	41840.52	-15915.50	-39.18
Work	0.82	0.83	-0.01	-2.06
Hours of Work	32.44	33.69	-1.25	-5.05
School Attendance	0.05	0.07	-0.02	-5.18
Male	0.54	0.50	0.04	6.12
Age of Entry in the US	8.77	19.96	-11.19	-154.34
Cognitive Difficulty	0.02	0.04	-0.02	-15.46
Ambulatory Difficulty	0.01	0.02	-0.01	-9.08
Mobility Difficulty	0.01	0.03	-0.02	-11.80
Self-Care Difficulty	0.01	0.01	-0.00	-4.47
Hearing/Seeing Difficulty	0.01	0.02	-0.01	-3.35
Observations	7,469	1,478,087		

Table A6: Descriptive Statistics; Sample: Entire Population (Bandwidth: 24 Quarters)

Variable	Mean		Difference	t-statistic
	Treatment Group	Control Group		
Treatment Group				
Total Income (in $1000)	25795.66	41613.18	-15817.52	-44.48
Work	0.82	0.83	-0.01	-2.67
Hours of Work	32.41	33.68	-1.27	-5.58
School Attendance	0.06	0.08	-0.02	-5.50
Male	0.54	0.50	0.04	6.83
Age of Entry in the US	8.76	19.94	-11.18	-170.20
Cognitive Difficulty	0.02	0.04	-0.02	-17.59
Ambulatory Difficulty	0.01	0.02	-0.01	-9.54
Mobility Difficulty	0.01	0.03	-0.02	-12.19
Self-Care Difficulty	0.01	0.01	-0.00	-5.11
Hearing/Seeing Difficulty	0.02	0.02	-0.00	-1.71
Observations	9,224	1,769,490		

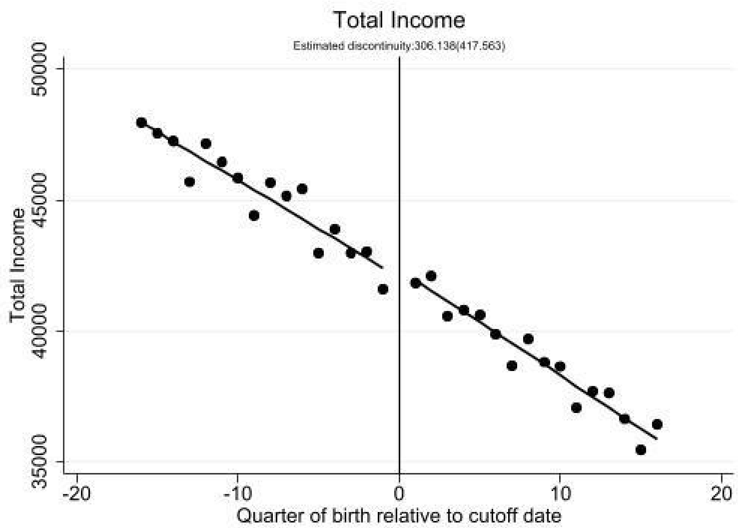


Figure A38: Total Income of the control group (used as robustness check) in the post DACA period (2014-2017) at a bandwidth of 16 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

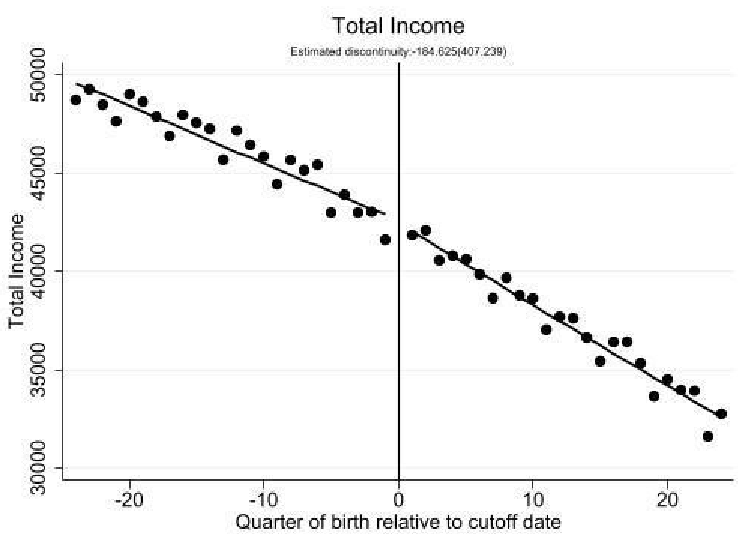


Figure A39: Total Income of the control group (used as robustness check) in the post DACA period (2014-2017) at a bandwidth of 24 quarters. The estimated discontinuity is computed using a linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

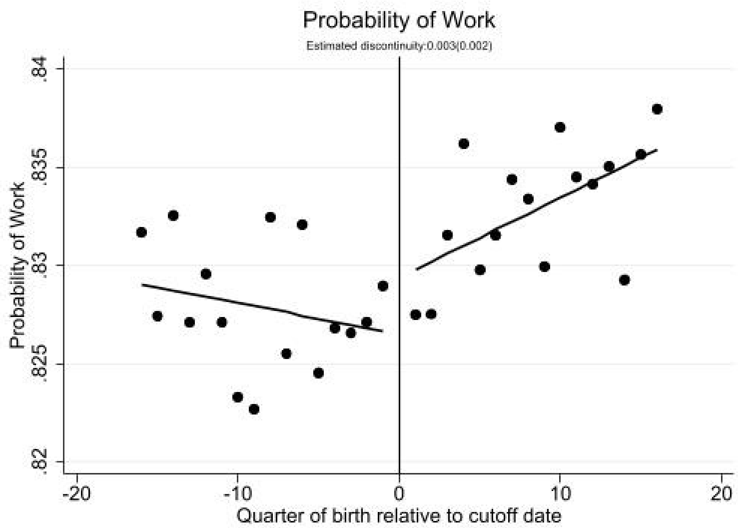


Figure A40: Probability of Work of the control group (used as robustness check) in the post DACA period (2014-2017) at a bandwidth of 16 quarters. The estimated discontinuity is computed using linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

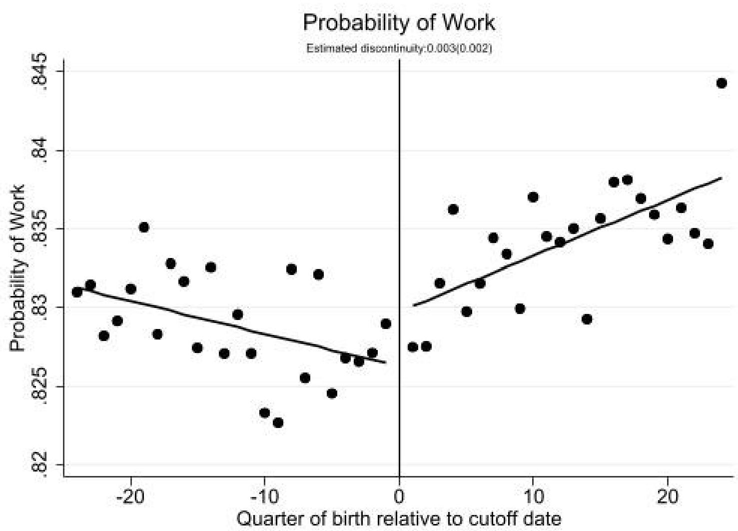


Figure A41: Probability of Work of the control group (used as robustness check) in the post DACA period (2014-2017) at a bandwidth of 24 quarters. The estimated discontinuity is computed using linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

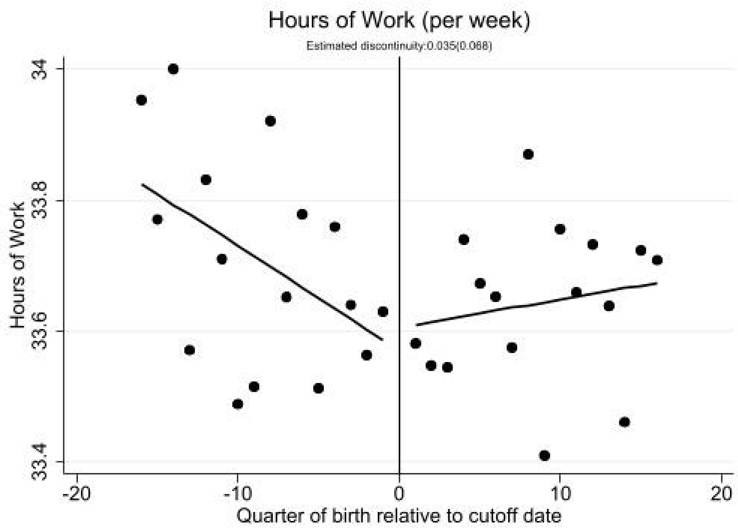


Figure A42: Hours of Work of the control group (used as robustness check) in the post DACA period (2014-2017) at a bandwidth of 16 quarters. The estimated discontinuity is computed using linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

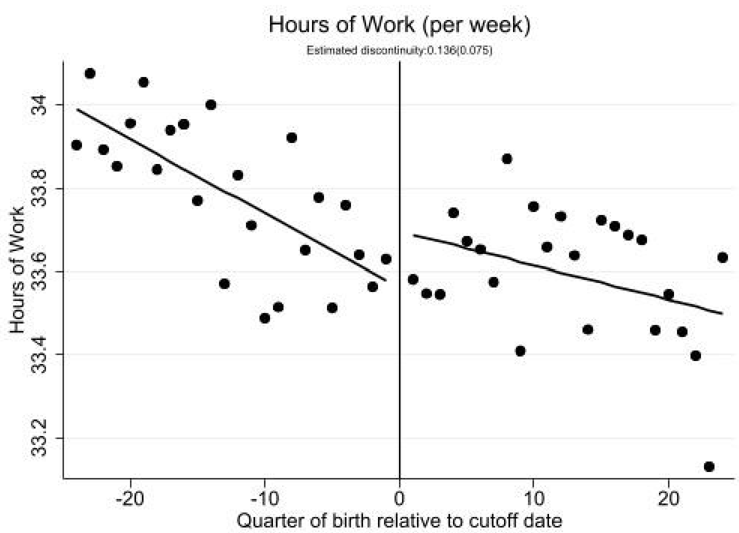


Figure A43: Hours of Work of the control group (used as robustness check) in the post DACA period (2014-2017) at a bandwidth of 24 quarters. The estimated discontinuity is computed using linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

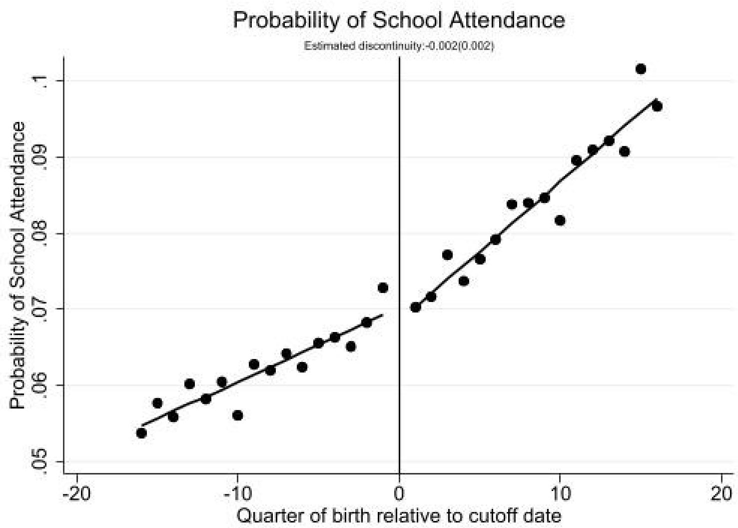


Figure A44: Probability of School Attendance of the control group (used as robustness check) in the post DACA period (2014-2017) at a bandwidth of 16 quarters. The estimated discontinuity is computed using linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

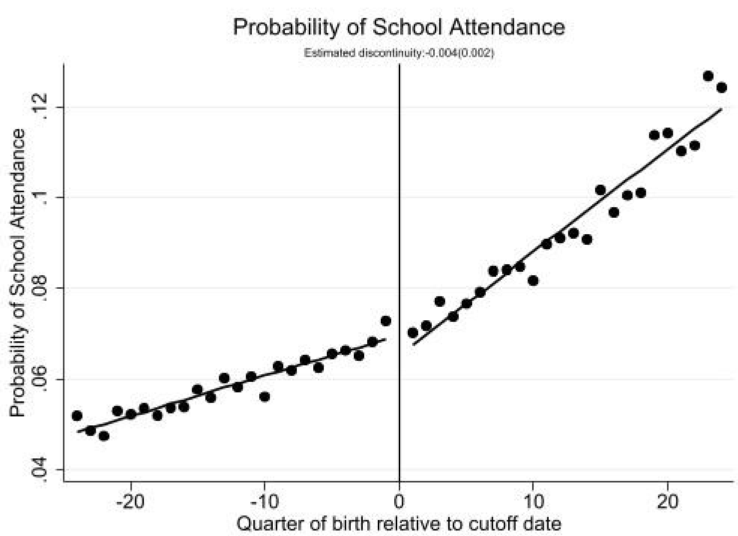


Figure A45: Probability of School Attendance of the control group (used as robustness check) in the post DACA period (2014-2017) at a bandwidth of 24 quarters. The estimated discontinuity is computed using linear regression; robust standard errors in parenthesis. The fitted values from this regression are plotted via lines.

Table A7: The Impact of DACA on Wage Income and Investment Income

Outcome Variable:	Wage Income						Investment Income					
	Baseline RD			DRD			Baseline RD			DRD		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Estimate	6819.975***	6508.506***	4463.382***	6304.425***	6330.839***	4267.546***	59.290	46.216	15.425	-15.177	-10.479	-7.836
	(2021.932)	(1733.256)	(1528.350)	(2064.951)	(1773.626)	(1565.399)	(47.814)	(38.266)	(32.119)	(71.164)	(53.702)	(52.160)
Bandwidth	16	20	24	16	20	24	16	20	24	16	20	24
Mean	23416.75	23435.27	23471.63	25582.51	25584.18	25534.87	39.006	51.219	52.171	149.941	139.348	145.066
Observations	5,731	7,469	9,224	130,449	162,946	195,628	5,731	7,469	9,224	130,449	162,946	195,628

Note: All estimates are weighted using sampling weights; Robust standard errors in parentheses The analysis is performed using (2) . * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A8: Results by Including Quarter Fixed Effects in the Baseline RD

Outcome Variable:	Total Income			Work or not			Hours of Work			School Attendance		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Estimate	6197.519***	6025.757***	4345.643***	0.028	0.034	0.028	2.240*	2.474**	1.688*	0.013	0.007	0.008
	(1997.362)	(1777.562)	(1534.186)	(0.027)	(0.023)	(0.021)	(1.257)	(1.096)	(1.004)	(0.014)	(0.012)	(0.011)
Bandwidth	16	20	24	16	20	24	16	20	24	16	20	24
Mean	25,650.63	25,925.02	25,795.67	0.817	0.820	0.819	32.419	32.442	32.418	0.057	0.059	0.061
Observations	5,731	7,469	9,224	5,731	7,469	9,224	5,731	7,469	9,224	5,731	7,469	9,224

Note: All estimates are weighted using sampling weights; Robust standard errors in parentheses The analysis is performed using (2) . * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A9: Descriptive Statistics; Sample: Mexican Non-Citizens with a High School Degree (Bandwidth: 16 Quarters)

Variable	Mean		Difference	t-statistic
	Treatment Group	Control Group		
Treatment Group				
Total Income (in $1000)	25650.63	22295.95	3354.68	6.80
Work	0.82	0.76	0.06	7.71
Hours of Work	32.42	30.18	2.24	6.67
School Attendance	0.06	0.04	0.02	5.74
Male	0.54	0.54	0.00	0.10
Age of Entry in the US	8.81	22.61	-13.80	-149.38
Cognitive Difficulty	0.02	0.01	0.01	4.37
Ambulatory Difficulty	0.01	0.01	0.00	0.83
Mobility Difficulty	0.01	0.01	0.00	2.71
Self-Care Difficulty	0.00	0.00	0.00	2.09
Hearing/Seeing Difficulty	0.01	0.01	0.00	1.89
Observations	5,731	15,946		

Table A10: Descriptive Statistics; Sample: Mexican Non-Citizens with a High School Degree (Bandwidth: 16 Quarters)

Variable	Mean		Difference	t-statistic
	Treatment Group	Control Group		
Treatment Group				
Total Income (in $1000)	25925.02	22288.68	3636.34	7.97
Work	0.82	0.76	0.06	8.87
Hours of Work	32.44	30.22	2.22	7.55
School Attendance	0.06	0.04	0.02	7.17
Male	0.54	0.54	0.00	0.22
Age of Entry in the US	8.77	22.66	-13.89	-170.14
Cognitive Difficulty	0.02	0.01	0.01	4.53
Ambulatory Difficulty	0.01	0.01	0.00	1.48
Mobility Difficulty	0.01	0.01	0.00	3.15
Self-Care Difficulty	0.00	0.00	0.00	2.23
Hearing/Seeing Difficulty	0.01	0.01	0.00	1.75
Observations	7,469	19,707		

Table A11: Descriptive Statistics; Sample: Mexican Non-Citizens with a High School Degree (Bandwidth: 24 Quarters)

Variable	Mean		Difference	t-statistic
	Treatment Group	Control Group		
Treatment Group				
Total Income (in $1000)	25795.66	22431.43	3364.23	8.31
Work	0.82	0.77	0.05	9.26
Hours of Work	32.42	30.33	2.09	7.73
School Attendance	0.07	0.04	0.03	8.35
Male	0.54	0.54	0.00	0.25
Age of Entry in the US	8.76	22.75	-13.99	-187.19
Cognitive Difficulty	0.02	0.01	0.01	5.14
Ambulatory Difficulty	0.01	0.01	0.00	1.98
Mobility Difficulty	0.02	0.01	0.01	4.06
Self-Care Difficulty	0.00	0.00	0.00	2.54
Hearing/Seeing Difficulty	0.02	0.01	0.01	3.15
Observations	9,224	23,353		

Table A12: DRD results using Alternative Control Group; Sample: Mexican Non-Citizens with a High School Degree

Outcome Variable:	Total Income			Work or not			Hours of Work			School Attendance		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Estimate	5718.633**	5992.774***	3769.023**	0.014	0.021	0.012	1.583	1.870	0.793	0.018	0.007	0.011
	(2221.228)	(1972.392)	(1714.912)	(0.031)	(0.028)	(0.025)	(1.456)	(1.278)	(1.170)	(0.015)	(0.013)	(0.012)
Bandwidth	16	20	24	16	20	24	16	20	24	16	20	24
Mean	23,140.64	23,239.75	23,345.94	0.775	0.778	0.780	30.741	30.799	30.897	0.041	0.041	0.043
Observations	21,677	27,176	32,577	21,677	27,176	32,577	21,677	27,176	32,577	21,677	27,176	32,577

Note: All estimates are weighted using sampling weights; Robust standard errors in parentheses The analysis is performed using (2) . * p < 0.1, ** p < 0.05, *** p < 0.01

Table A13: Descriptive Statistics; Sample: Mexican Non-Citizens with a High School Degree Satisfying the Year of Immigration Requirements (Bandwidth: 16 Quarters)

Variable	Mean		Difference	t-statistic
	Treatment Group	Control Group		
Treatment Group				
Total Income (in $1000)	25650.63	21967.92	3682.71	7.28
Work	0.82	0.77	0.05	6.61
Hours of Work	32.42	30.38	2.04	5.82
School Attendance	0.05	0.03	0.02	5.90
Male	0.55	0.55	-0.00	-0.30
Age of Entry in the US	8.81	20.54	-11.73	-136.62
Cognitive Difficulty	0.02	0.01	0.01	4.52
Ambulatory Difficulty	0.01	0.01	0.00	0.41
Mobility Difficulty	0.01	0.00	0.01	2.98
Self-Care Difficulty	0.00	0.00	0.00	2.13
Hearing/Seeing Difficulty	0.01	0.01	0.00	1.99
Observations	5,731	12,054		

Table A14: Descriptive Statistics; Sample: Mexican Non-Citizens with a High School Degree Satisfying the Year of Immigration Requirements (Bandwidth: 20 Quarters)

Variable	Mean		Difference	t-statistic
	Treatment Group	Control Group		
Treatment Group				
Total Income (in $1000)	25925.03	21995.5	3929.53	8.37
Work	0.82	0.77	0.05	7.56
Hours of Work	32.44	30.43	2.01	6.52
School Attendance	0.06	0.03	0.03	7.64
Male	0.55	0.55	-0.00	-0.19
Age of Entry in the US	8.78	20.62	-11.84	-155.78
Cognitive Difficulty	0.02	0.01	0.01	4.61
Ambulatory Difficulty	0.01	0.01	0.00	0.92
Mobility Difficulty	0.01	0.00	0.01	3.42
Self-Care Difficulty	0.00	0.00	0.00	2.14
Hearing/Seeing Difficulty	0.01	0.01	0.00	1.96
Observations	7,469	14,785		

Table A15: Descriptive Statistics Sample: Mexican Non-Citizens with a High School Degree Satisfying the Year of Immigration Requirements (Bandwidth: 24 Quarters)

Variable	Mean		Difference	t-statistic
	Treatment Group	Control Group		
Treatment Group				
Total Income (in $1000)	25795.66	22072.46	3723.20	8.94
Work	0.82	0.77	0.05	7.55
Hours of Work	32.42	30.58	1.84	6.50
School Attendance	0.06	0.03	0.03	8.91
Male	0.55	0.55	-0.00	-0.18
Age of Entry in the US	8.76	20.72	-11.96	-171.75
Cognitive Difficulty	0.02	0.01	0.01	5.20
Ambulatory Difficulty	0.01	0.01	0.00	1.47
Mobility Difficulty	0.01	0.00	0.01	4.27
Self-Care Difficulty	0.00	0.00	0.00	2.47
Hearing/Seeing Difficulty	0.02	0.01	0.01	3.34
Observations	9,224	17,375		

Table A16: DRD Results using Alternative Control Group: Sample: Mexican Non-Citizens with a High School Degree Satisfying the Year of Immigration Requirements

Outcome Variable:	Total Income			Work or not			Hours of Work			School Attendance		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Estimate	5224.473**	5621.905***	3299.667*	0.002	0.004	-0.000	1.123	1.224	0.228	0.020	0.008	0.011
	(2255.449)	(2020.641)	(1751.124)	(0.033)	(0.029)	(0.026)	(1.508)	(1.330)	(1.216)	(0.015)	(0.014)	(0.013)
Bandwidth	16	20	24	16	20	24	16	20	24	16	20	24
Mean	23,091.58	23,243.48	23,306.38	0.782	0.786	0.788	31.005	31.070	31.190	0.041	0.041	0.042
Observations	17,785	22,254	26,599	17,785	22,254	26,599	17,785	22,254	26,599	17,785	22,254	26,599

Note: All estimates are weighted using sampling weights; Robust standard errors in parentheses The analysis is performed using (2) . * p < 0.1, ** p < 0.05, *** p < 0.01

Appendix B

Chapter 2

Appendix:

Table B1: Percentage of days a particular food item was offered at Crossroads during the study window

Protein Foods

Food Item	Food Item Size	Points Required for Purchase	% of days offered during study window
Beef Stew - USDA	24 oz bag	24	58.79
Eggs - Whole Fresh - Small	1 doz	12	50.75
Peanut Butter (Algood)	18oz jar	18	50.25
Sliced Smoked Turkey Breast	2 lbs	28	45.73
Ensure Nutritional Shake	8 oz	2	25.13
Cheese (Mild Cheddar Shredded)	5 lbs	20	24.12
Cashews - Maple Bacon flavored	8 oz bag	4	23.62
Peanut Butter (SK)	16oz	16	15.08
Frozen - Classic Mac & Cheese	30 oz	30	14.57
Tuna Starkist - canned	12 oz can	7	14.57
Beef Knockwurst - Sausage	10 oz	10	12.06
Eggs - Whole Fresh USDA Eggs	1 doz	12	11.56
Peanuts - Roasted in shell	16 oz	16	11.56
Cashews - Maple Bacon Flavored	8 oz	4	10.05
Peanut Butter Assorted	16 oz	16	9.55
Pepperoni Slices	10 lb box	160	8.54
Eggs Whole Eggs Liquid in Carton	32 oz	24	8.54
USDA - Catfish Fillets	2 lbs	32	8.54
Cheese Mild Cheddar - Shredded 2 lbs	2 lbs	20	8.54
Beef - Ground Meat	16 oz	12	8.04
Chicken Leg Quarters (TB)	5 lb	50	7.54
Beef Sausage 100%Grass-Fed Beef	12 oz package	12	7.54
Ham, cooked and frozen	3 lb box	39	7.04
Yogurt - Single Cup	5 oz	1	7.04
Burritos - Grilled Chicken 4 pack	40 oz	20	6.53
Tacos - Sausage, Egg and Cheese	1.8 lbs	20	6.53
Shrimp Taco Bowl	10 oz	6	6.03
Lunchables - Pepperoni Pizza	4.3 oz	3	6.03
XL Beef, Bean & Cheese Burrito	12 oz	6	5.53
Breakfast Sandwich - Bacon Cheddar	5 oz	2	5.53

Chicken Breast Strips - Fully Cooked	5 lbs	80	5.53
Beef - Chopped Philly Steak - Raw Frozen	10 lbs	160	5.53
Eggs Liquid Egg Whites	16 oz	12	5.03
Chicken Spicy Breaded Breast Strips	21 oz	21	4.52
Beef Junior Franks (little smokies) Grass-Fed Beef	10 oz package	10	4.52
Tamales - Pork	48 oz	30	4.02
Pork Sausage - Frozen	1 lb	10	4.02
Eggs (MF) Large	dozen	12	4.02
Chicken - Boneless skinless thighs	2.32 Lbs	32	4.02
Jambalaya Red Lentil	2 oz	2	3.52
Beef Fajitas Fully Cooked	14 oz	14	3.02
Eggs (MF) Medium	1 doz carton	12	3.02
Beef Patties - with Jalapeno and Pepper Jack	1.33 lbs	22	3.02
Turkey, Whole 12.5 lbs average	12.5 lbs.	90	3.02
Chicken Whole	6 lbs	44	2.51
Tuna	5 oz can	4	2.51
Chicken Whole - USDA	4 lb pkg	30	2.51
Chicken Organic boneless skinless leg & thighs	6 lbs	96	2.51
Chicken Breast - Boneless Skinless	72 oz	72	2.51
Beef Fajitas - Seasoned, Frozen, Raw	5 lbs	80	2.51
Cauliflower Curry Power Bowls	10 oz	10	2.51
Turkey Whole Cooked	12.5	100	2.51
Chicken Taquitos	10 oz	8	2.01
Turkey Breast - Cajun Style - Cooked	4.75 lbs	70	2.01
Chicken Taquitos	4.5 lbs	58	2.01
Rockfish Bowl	10 oz	6	2.01
Ground Beef 93% Lean	1 lb	12	1.51
Chicken Drmsticks	5 lbs	60	1.51
Chicken Breast - Boneless Skinless (Small)	2 lbs	28	1.51
Burrito - Beef, Bean and Cheese	14 oz	7	1.51
Soups, assorted, canned	12 oz cans	1	1.01
Sausage and Red Beans	5 lbs	45	0.50
Nuts, salted and roasted mix	44 oz	25	0.50

Dairy Products

Food Item	Food Item Size	Points Required for Purchase	% of days offered during study window
Dairy - 1% Natrel Shelf Stable	32 oz	1	62.31
Evaporated Dairy - Can	12 oz	2	15.58
USDA Whole Dairy	64 oz	2	9.55
Yogurt 8 pouches	28 oz	5	8.04
Half and Half Organic	32	8	7.04
Yougurt - Gogurt	32 oz	8	6.03

Cheese Mozzarella Shredded	1 lb	4	5.53
Almond Dairy Original (Silk)	64 oz	2	4.52
Cream - Heavy Whipping	16 oz	8	4.02
1 % Lowfat Dairy Parmalat	32oz box	1	3.52
2% Fresh Organic Dairy	64oz	2	2.51

Vegetables

Food Item	**Food Item Size**	**Points Required for Purchase**	**% of days offered during study window**
Green Beans - Signature	14.5 can	1	74.87
Corn (Signature)	15.25 oz can	1	73.37
Beans Dry - Dark Red Kidney Beans	2 lb bag	6	69.35
Onions - Fresh	4lb Portion	4	65.33
Pinto Beans - Dry - 32 oz bag	2 lb bag	6	55.78
Carrots - fresh	4 lb portion	4	52.26
Potatoes- Fresh	5lb bag	5	48.24
Vegetarian Beans (MM)	15 oz can	1	45.23
Cabbage - Fresh	4 lbs	4	44.22
Pinto Beans (SK)	2lb bag	6	42.21
Vegetarian Beans (Lakeside)	16 oz can	1	42.21
Tomato Sauce (DM)	15 oz can	1	32.66
Squash, Spaghetti, Fresh	4lb	4	32.16
Vegetarian Soup (H)	10.75 oz can	1	30.65
Bell Peppers - Fresh	4 lb portion	4	30.65
Sweet Potatoes - Fresh	4 lb portion	4	30.15
Refried Beans (Hart Brand)	16oz can	1	26.13
Tomatoes - fresh	4 lbs portion	4	25.63
Salsa Chunky Mild - Pace	64 oz jar	8	24.62
Squash, Yellow, Fresh	4 lb portion	4	21.61
Salsa - Organic Medium Pace	24 oz	4	21.61
Potatoes - fresh	10lb	10	20.10
Spaghetti Sauce USDA	15 oz	2	20.10
Refried Beans (TF)	15.25 oz can	1	19.60
Potatoes - Sliced Canned	14.5	1	18.59
Cucumbers - Fresh	4 lb bag	4	18.09
Lentils	32 oz	6	16.58
Zucchini, Fresh	4 lb portion	4	15.58
Butternut Squash - Fresh	4lb portion	4	15.08
Tomatoes Cherry	11 oz package	1	14.07
Tomatoes - Cherry	12 oz container	1	14.07
Beans - Light Red Kidney Beans DRY	2 lb bag	6	13.57

Carrots - fresh	1 lb bag	1	11.06
Corn, Fresh	4 lb portion	4	10.55
Split Peas - Yellow	32 oz	6	10.55
Brussels Sprouts	2 lb bag	2	10.05
Carrots, canned (Hart)	14.5 oz can	1	9.55
Beans, Borlotti	13 oz box	1	8.54
Peas, Carots & Beets Baby Food	21 oz	12	8.54
Mushroom Fresh	8 oz pkg	1	8.54
Baby Bok Choy	4 lbs	4	8.04
Broccoli - Fresh	4 lbs	4	8.04
Corn (DM)	15.25 oz can	1	7.54
Bell Peppers Mini Peppers - Fresh	1 lb portion	1	7.54
Turnips - Fresh	4 lbs	4	7.54
Eggplant, fresh	4lb	4	7.54
Green Beans PouchHello Fresh	6 oz package	1	7.54
Diced Tomatoes (Vine Ripe)	14.5 oz can	1	7.54
Pinto Beans (K)	15 oz can	1	6.53
Acorn Squash	1 squash	4	6.53
Kale Greens	16 oz	1	6.03
Tomatoes, diced (365)	28oz	2	6.03
Sweet Peas (Del Monte) - Can	15 oz can	1	6.03
Celery - Fresh	4 lb portion	4	6.03
Green Beans (SK)	14.5 oz can	1	5.53
Tomatoes Vine Ripened Fresh	1 lb pkg	1	5.53
spinach, baby	5 oz bag	1	5.53
Lettuce, Romaine	Multi Pack	4	5.03
Carrots, Diced, Frozen	2.5lbs	4	4.52
Celery - Fresh	4lb portion	4	4.52
Red Potatoes, fresh	5 lb bag	5	4.52
Tomatoes, Sweet Yellow Fresh	10.5 oz bowl	1	4.02
Cabbage - Red sliced	8 oz	1	4.02
Onions Purple	4 lb portion	4	4.02
Salad - Asian Chopped Kit	13 oz	2	4.02
Greens - Mixed	5 oz bag	1	3.52
Beets - Fresh	4 lb bag	4	3.52
Cauliflower fresh	4 lbs	4	3.52
Potatoes 8lb -fresh	8lb bag	8	3.52
Lettuce - Spring Mix - Organic	5 oz container	1	3.02
Jalapeno Peppers Fresh	4 lbs	4	3.02
Cauliflower Florets	10 oz	1	3.02
Green Beans - fresh - Steam Bag	8 oz	1	3.02
Spinach Bowl	11 oz	1	3.02
Leeks - Fresh	4 lbs	4	3.02
Onions - Green	2 Lbs	2	2.51
Green Beans - fresh	4 lb portion	4	2.51
Lettuce - 1 head	Medium	1	2.51

Salad - Ceasar with Chicken	12.5 oz	4	2.01
Salad Bowl - Ceasar with Chicken	12 oz	4	2.01
Spinach Fresh	2.5 lb bag	2	2.01
Radish, Fresh	4lbs	4	2.01
Snow Peas	8 oz	1	1.51
Holiday Dinner Kit - Frozen	7 lbs	30	1.51
Vegetables Assorted Tray	4 lb	10	1.51
Arugula	4 small bags	1	1.51
Poblano Peppers - fresh (1 pkg.)	4 lbs	4	1.51
Broccoli Florets	2 lbs	2	1.01
Carrots - Fresh - Rainbow	12 oz bag	1	1.01
Tomatillos (Green Tomatoes) - Fresh	4 lb portion	4	1.01
Pumpkin	10 lbs	10	1.01
Squash, Summer - Fresh	4 lbs	4	1.01
Broccoli Cauliflower Mix	12 oz	1	1.01
Cherry Tomatoes (Assorted)	2 lbs	2	1.01
Mix Sweet Peppers	2 lb	2	0.50
Cucumbers 50 lb box	4 lbs	4	0.50
Spinach - Fresh - Bagged	2 lb	2	0.50
Greens,Turnip - Fresh	4 lb portion	4	0.50

Grains

Food Item	**Food Item Size**	**Points Required for Purchase**	**% of days offered during study window**
Spaghetti (M)	16oz box	16	78.89
Cereal - Bran Flakes (RF)	17.3 oz box	17	60.80
Oatmeal - Quick Cooking Rolled Oats	3 lb bag	48	51.76
Cereal - Corn Biscuits (RF)	14 oz box	14	45.23
Macaroni & Cheese Dinner	7.25 oz box	7	39.20
Rice 2 lb large bag	32 oz bag	32	38.69
Elbow Macaroni (M)	16oz	16	32.66
Pasta - Elbow Macaroni	16 oz	16	29.65
Cereal - Tasteeos (RF)	14 oz box	14	29.15
Cereal Cream of Wheat / Farina	18 oz	18	28.14
Rice 2 lb (PE30)	32 oz bag	32	27.64
Rice, Long Grain (Signature Kitchens)	32 oz bag	32	20.60
Cereal - Crispy Rice	12 oz box	12	19.60
Quinoa - Heat and Serve Cups	8.8 oz	8	18.59
Cereal - Crispy Rice (RF)	12 oz box	12	18.09
Bagels Plain - half dozen	17 oz	17	11.06
Pasta - Fideo / Vermicelli bag	7 oz	7	10.05
Cereal - Rice Crispies (RF)	12 oz box	12	9.05
Classic Mac & Cheese Frozen	30 oz	30	8.04
Rice - Fried - Ready to Serve	9 oz	8	7.04

Oatmeal - Quick Oats	42 oz	42	6.53
Cereal - Corn Flakes (RF)	18oz	18	6.03
Nutri-Grain Bars Pumpkin Spice	1 box	10	5.53
Pasta - Elbow Macaroni Whole Grain	16 oz	16	5.53
Oatmeal - Old Fashioned	18 oz	18	4.52
Cereal - Lucky Charms Bowl	1 oz	1	3.52
Bagels - Frozen large bag	36 oz	36	3.52
Breaded Onion Rings	2 lbs	20	3.02
Rice heat and serve cups	9	3	2.51
Mac and Cheese - Microwave Cups	8.2 ox	9	2.51
Cereal - Cocoa Puffs Bowl	1 oz	1	2.01
Cereal - Multigrain Oats and Quinoa	16 oz	16	2.01
Brown Rice - Family Bowl Ready to Serve	16 oz	16	1.51
Spaghetti, Whole Wheat (PS)	32 oz.	32	1.01

Fruits

Food Item	Food Item Size	Points Required for Purchase	% of days offered during study window
Blueberry -frozen 3 Lb bag	3lb bag	6	74.37
Figs - Dry	1 lb bags	3	65.83
Plums - Dry Pitted	1 lb bags	3	54.77
Watermelon - fresh	10lb	10	38.69
Grape Juice	64 oz bottle	4	30.65
Grapefruit - Fresh	4 lb Portion	4	21.11
Raisins, 6 pack (Sun Maid)	8 oz	1	20.10
Strawberries Frozen Cup	4.5 oz	1	16.58
Olives - Organic Pimento Stuffed	10 oz	4	16.08
Peaches (DM) - canned	15 oz can	1	15.08
Pear Halves	15 oz can	1	15.08
Apple Sauce Cups (White House) x4	4 cups	1	14.07
Apple Fruit Crisp	.35 oz	1	14.07
Pears - fresh	4lb portion	4	13.07
Apples - Fresh	4 lb portion	4	12.06
Blueberries-Frozen	2.5lbs	5	11.06
Cherries, Dried (Shoreline)	2lb bag	8	10.55
Grapes, fresh	1 lb bag	1	10.05
Honey Dew Melon- Fresh	5 lb melon	5	10.05
Pineapple-fresh	1 medium pineapple	4	9.55
Mango, Fresh	4 lb bag	4	8.54
Banana Blueberry Gerber Grabbers	16 oz	8	8.54
Oranges - Fresh	4lb portion	4	8.54
Limes	2 lb bag	2	8.04

Cantaloupe - Fresh	1 melon	4	8.04
Oranges - Fresh	4 lb portion	4	7.54
Raisins (LB)	15 oz box	1	7.54
Bananas - Fresh	4 lbs	4	7.04
Watermelon Small - Fresh	5 lb	5	6.53
Apple Cinnamon Bites - Snack	8 oz	2	5.53
Papaya - fresh	4 lb	4	5.53
Strawberries - fresh	1 lb tray	1	4.52
Avocados, fresh	4lb bag	4	4.02
Peaches, Fresh	4 lb	5	3.52
Plums - Fresh	4 lbs	4	3.02
Juice - Passion Fruit 100% Juice	144 oz	9	2.51
Mixed Fruit Frozen	2 lbs	4	2.51
Blackberries - Fresh	12 oz	1	2.51
Plantains	5 lb portion	5	2.51
Cranberries - Fresh	1 lb	1	2.01
Blueberries Fresh	24 oz	1	2.01
Lemons	4lb portion	4	1.51
Red Raspberries	6 oz	1	1.01
Cherries Fresh	2 lbs	2	0.50

Table B2: Summary Statistics for Covariates and Outcomes for Households in the Balanced Diet Regime Period (N=560 households and 2571 observations*)

Variables		**Mean**	**Std. Dev.**	**Min**	**Max**
Covariates					
Household Income ($)	Overall	1176.262	756.668	0	5640
	Between		736.938		
	Within		313.165		
SNAP Amount	Overall	74.521	141.724	0	794
	Between		146.657		
	Within		42.285		
Household Size	Overall	4.464	2.564	1	12
	Between		2.528		
	Within		0.309		
Indicator of Kids in Household	Overall	0.638	0.481	0	1
	Between		0.480		
	Within		0.065		
Outcome Variables					
Percent- Protein	Overall	0.346	0.087	0	0.877
	Between		0.055		
	Within		0.072		
Percent-Vegetables	Overall	0.121	0.035	0	0.370
	Between		0.024		
	Within		0.028		
Percent- Dairy Products	Overall	0.006	0.006	0	0.046
	Between		0.005		
	Within		0.005		
Percent- Grains	Overall	0.470	0.085	0	0.832
	Between		0.057		
	Within		0.068		
Percent-Fruits	Overall	0.055	0.030	0	0.338
	Between		0.021		
	Within		0.025		
Percent-Miscellaneous	Overall	0.001	0.003	0	0.026
	Between		0.002		
	Within		0.000		
Number of Categories	Overall	4.940	0.709	2	6
	Between		0.651		
	Within		0.326		
Percent-Fresh Produce	Overall	0.101	0.039	0	0.365
	Between		0.028		
	Within		0.031		
Percent-Frozen Produce	Overall	0.006	0.010	0	0.102
	Between		0.007		
	Within		0.008		
Percent-Shelf Stable Produce	Overall	0.070	0.030	0	0.346
	Between		0.020		
	Within		0.025		

*On average, each household visited 4.59 times during the Balanced Diet Regime.

Table B3: Summary Statistics for Covariates and Outcomes for Households in the Open-Regime Period (N=560 households and 1498 observations*)

Variables		Mean	Std. Dev.	Min	Max
Covariates					
Household Income ($)	Overall	952.883	804.597	0	5640
	Between		688.280		
	Within		454.904		
SNAP Amount	Overall	76.856	137.809	0	760
	Between		139.752		
	Within		35.584		
Household Size	Overall	4.494	2.621	1	13
	Between		2.596		
	Within		0.215		
Indicator of Kids in Household	Overall	0.644	0.479	0	1
	Between		0.485		
	Within		0.018		
Outcome Variables					
Percent- Protein	Overall	0.437	0.183	0	1
	Between		0.134		
	Within		0.142		
Percent-Vegetables	Overall	0.189	0.101	0	0.718
	Between		0.075		
	Within		0.075		
Percent- Dairy Products	Overall	0.011	0.020	0	0.303
	Between		0.017		
	Within		0.015		
Percent- Grains	Overall	0.283	0.160	0	1
	Between		0.108		
	Within		0.131		
Percent-Fruits	Overall	0.079	0.048	0	0.545
	Between		0.037		
	Within		0.036		
Percent-Miscellaneous	Overall	0.001	0.002	0	0.018
	Between		0.003		
	Within		0.001		
Number of Categories	Overall	4.544	0.943	1	6
	Between		0.786		
	Within		0.568		
Percent-Fresh Produce	Overall	0.109	0.062	0	0.397
	Between		0.046		
	Within		0.047		
Percent-Frozen Produce	Overall	0.018	0.023	0	0.250
	Between		0.017		
	Within		0.018		
Percent-Shelf Stable Produce	Overall	0.141	0.085	0	0.636
	Between		0.063		
	Within		0.064		

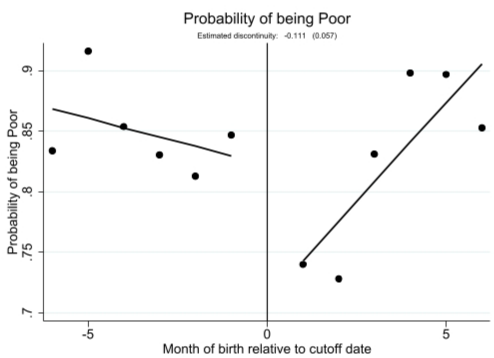
*On average, each household visited 2.68 times during the Open Regime.

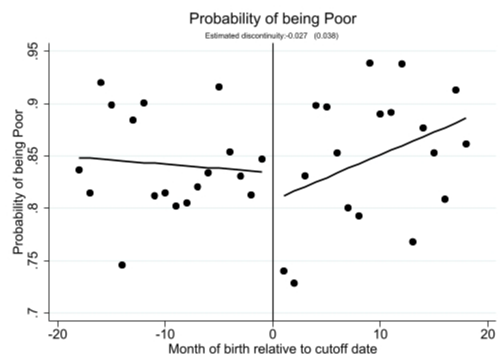
Appendix C

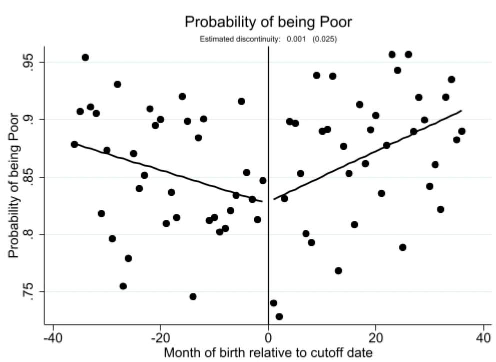
Chapter 3

Section 1: Continuity of Covariates (Outcome 1) [39]

Figure C1: Probability of being poor (Bandwidth Order: 6 months, 18 months, and 36 months)







[39] The estimated discontinuity is computed using linear regression ; robust standard errors in parenthesis; the fitted values from this regression are plotted via lines. Note: 95% Confidence Intervals not plotted due to the existence of state sampling weights in the data; Code for graphs has been taken from Clark and Royer (2013)

Figure C2: Probability of being Middle Income (Bandwidth Order: 6 months, 18 months, and 36 months)

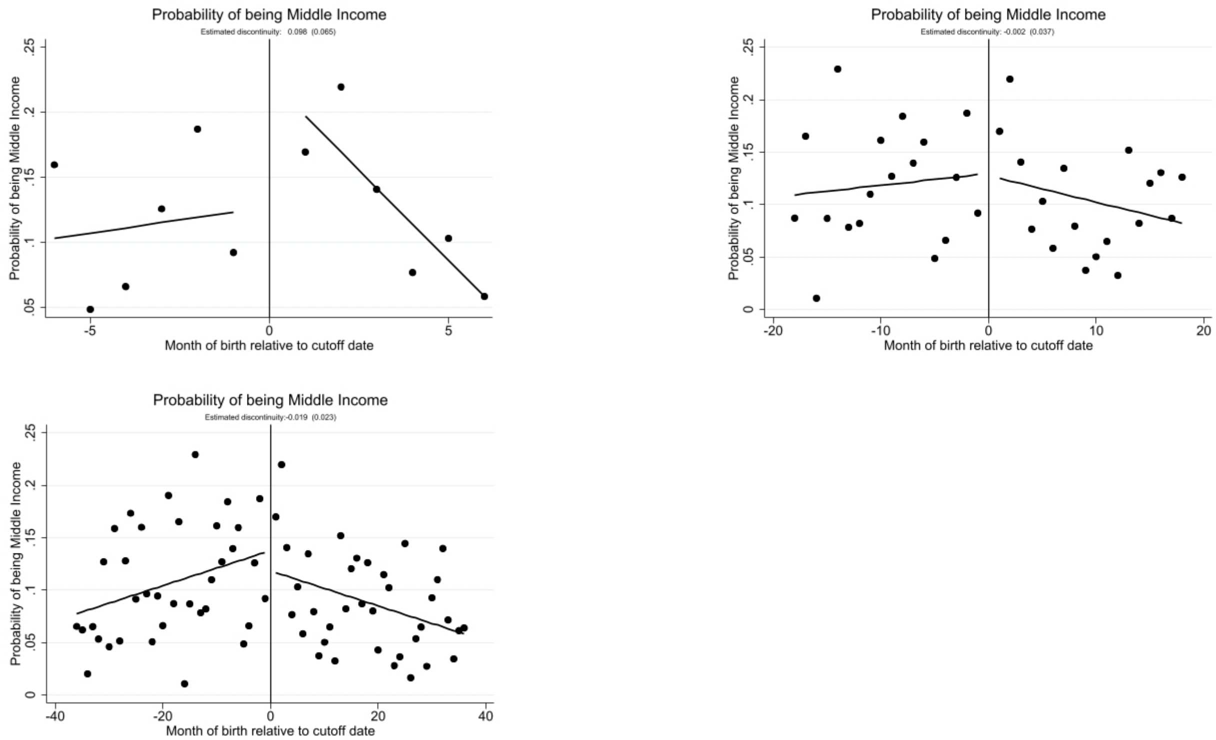


Figure C3: Probability of being Rich (Bandwidth Order: 6 months, 18 months, and 36 months)

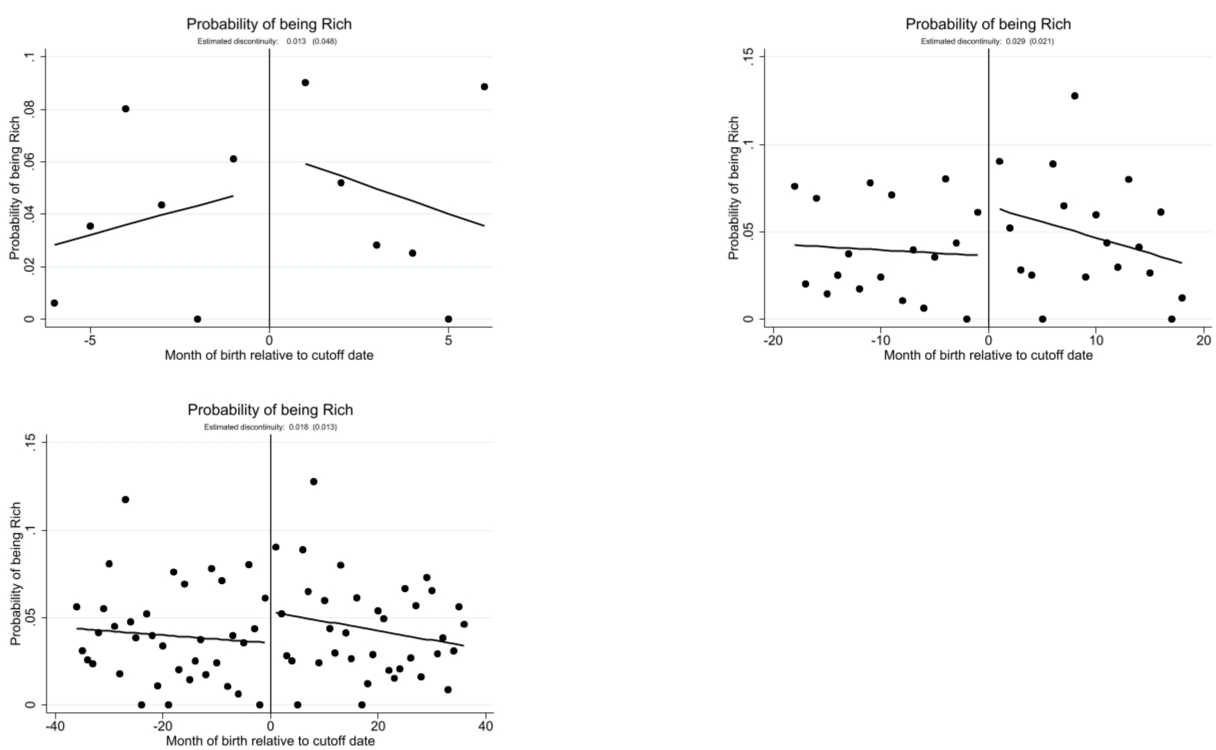


Figure C4: Probability of being a Hindu (Bandwidth Order: 6 months, 18 months, and 36 months)

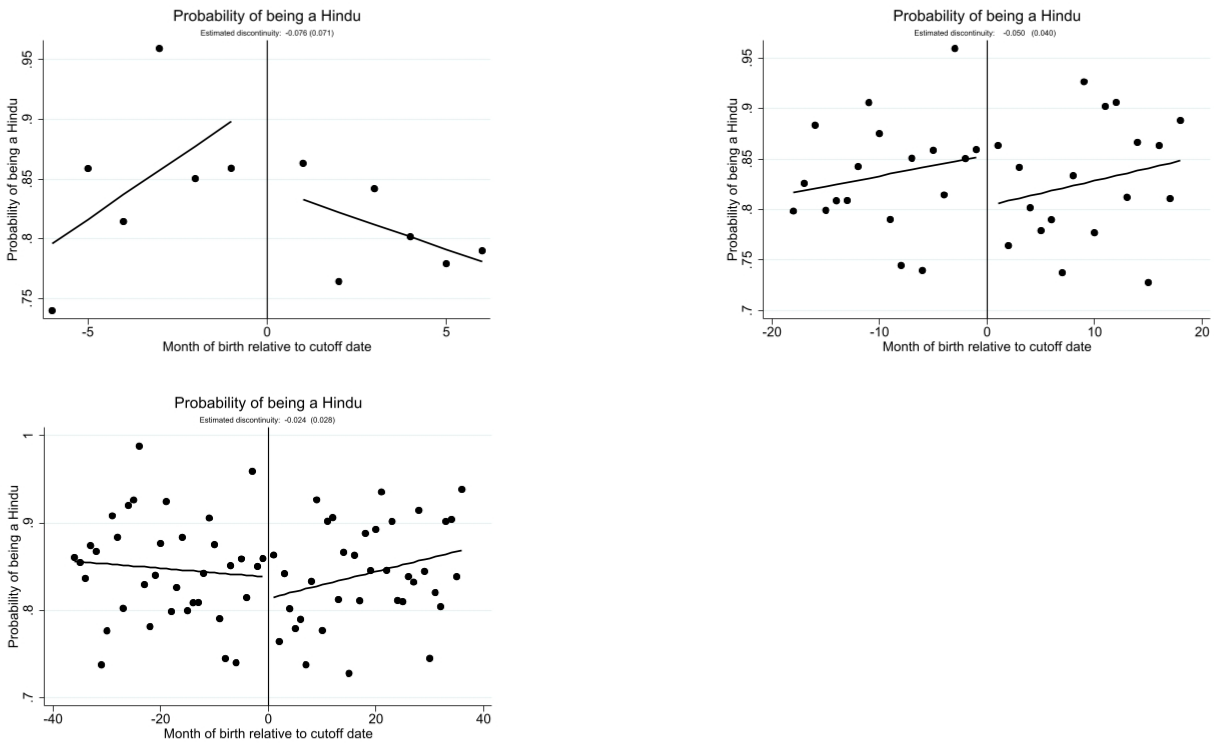


Figure C5: Probability of being a Muslim (Bandwidth Order: 6 months, 18 months, and 36 months)

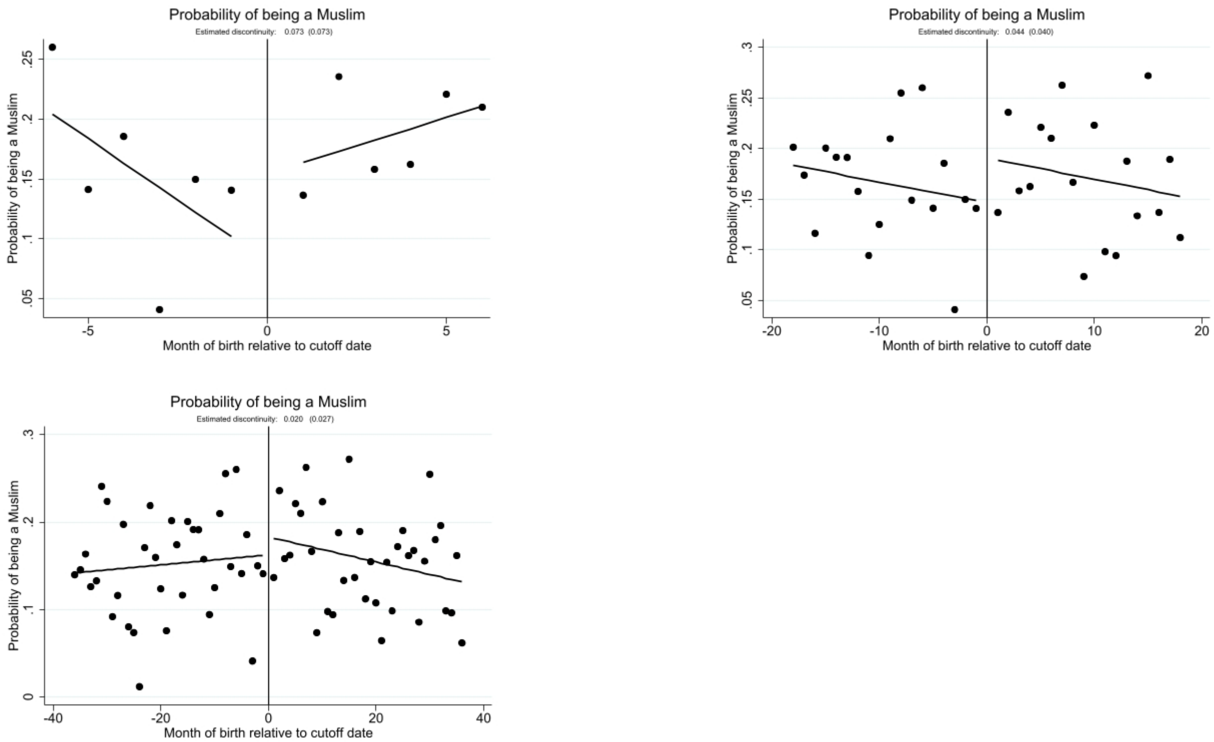


Figure C6: Probability of living in an urban area (Bandwidth Order: 6 months, 18 months, and 36 months)

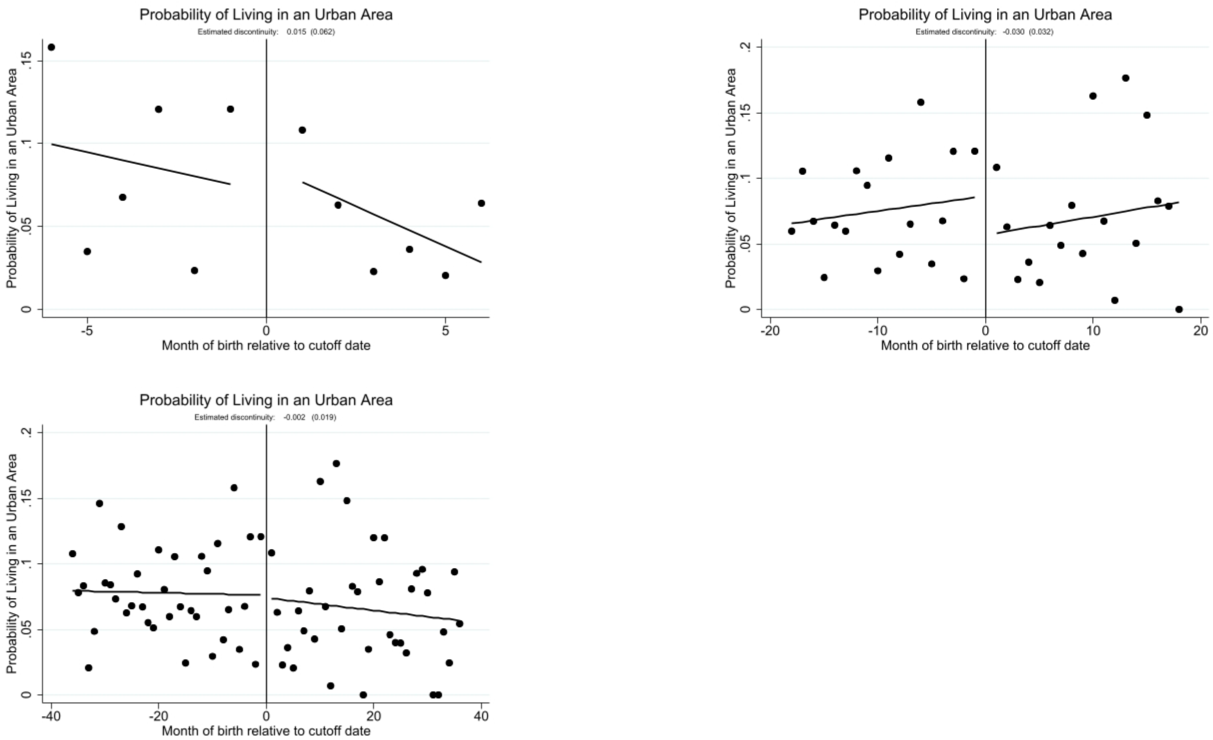
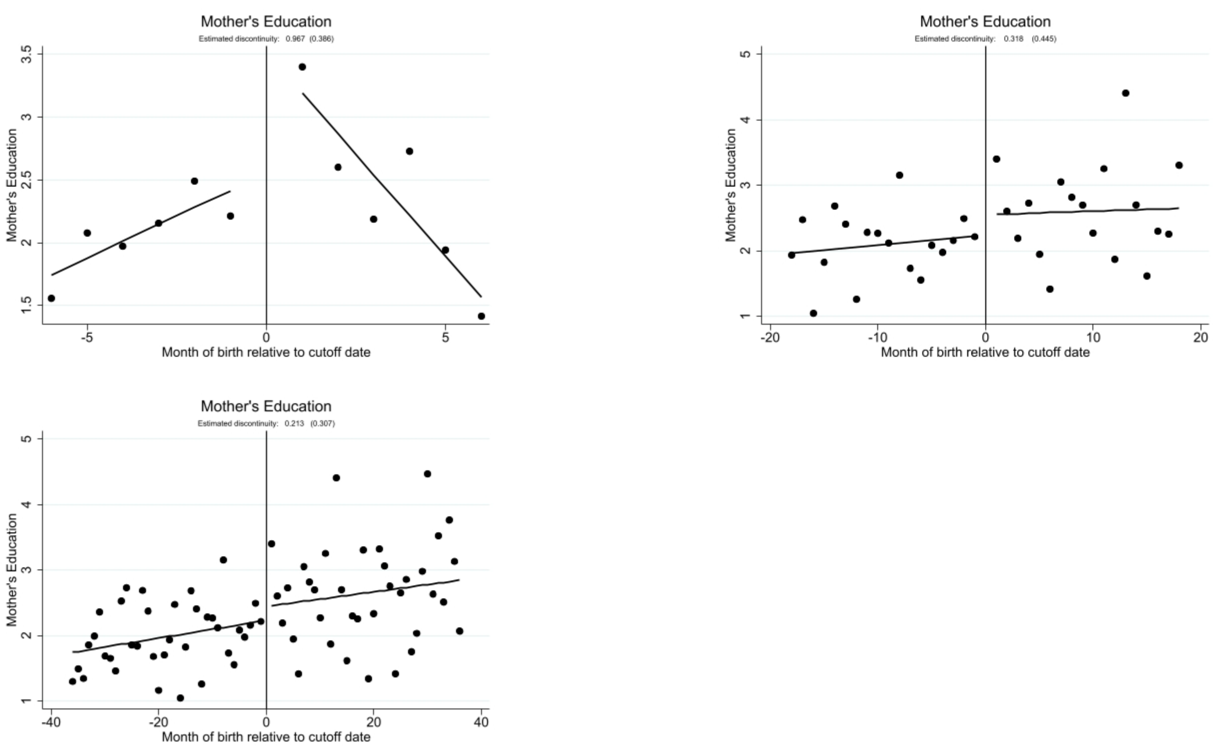


Figure C7: Mother's Education (Bandwidth Order: 6 months, 18 months, and 36 months)



Section 2: Continuity of Covariates (Outcome 2) [40]

Figure C8: Probability of being poor (Bandwidth Order: 6 months and 12 months)

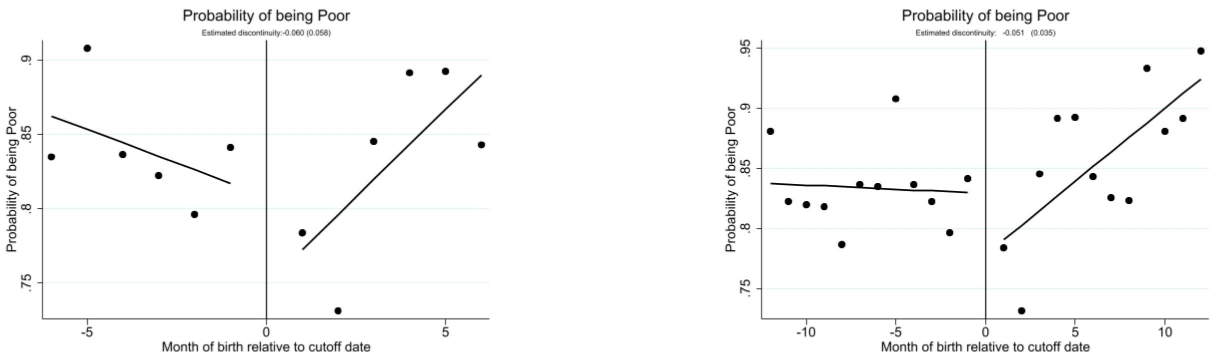


Figure C9: Probability of being Middle Income (Bandwidth Order: 6 months and 12 months)

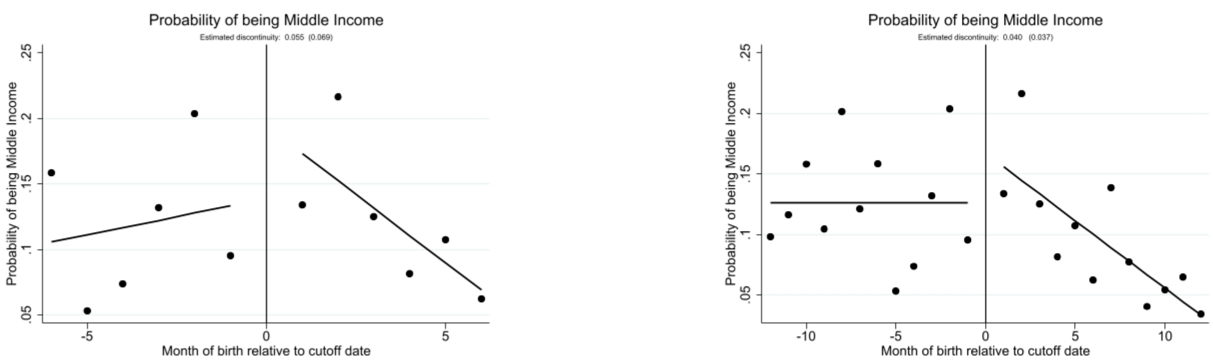
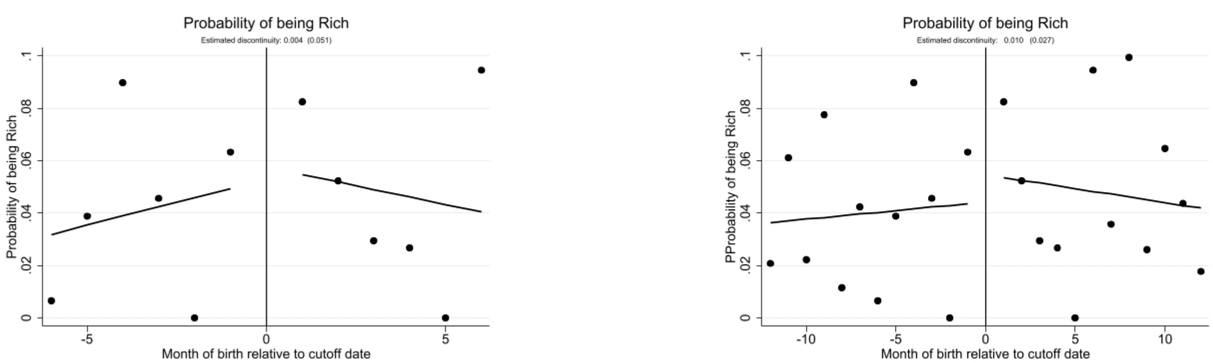


Figure C10: Probability of being Rich (Bandwidth Order: 6 months and 12 months)



[40] The estimated discontinuity is computed using linear regression ; robust standard errors in parenthesis; the fitted values from this regression are plotted via lines. Note: 95% Confidence Intervals not plotted due to the existence of state sampling weights in the data; Code for graphs has been taken from Clark and Royer (2013)

Figure C11: Probability of being a Hindu (Bandwidth Order: 6 months and 12 months)

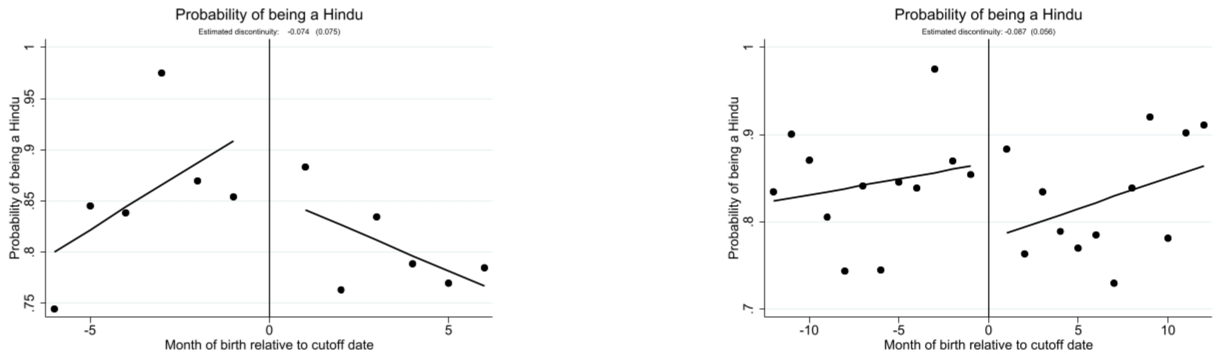


Figure C12: Probability of being a Muslim (Bandwidth Order: 6 months and 12 months)

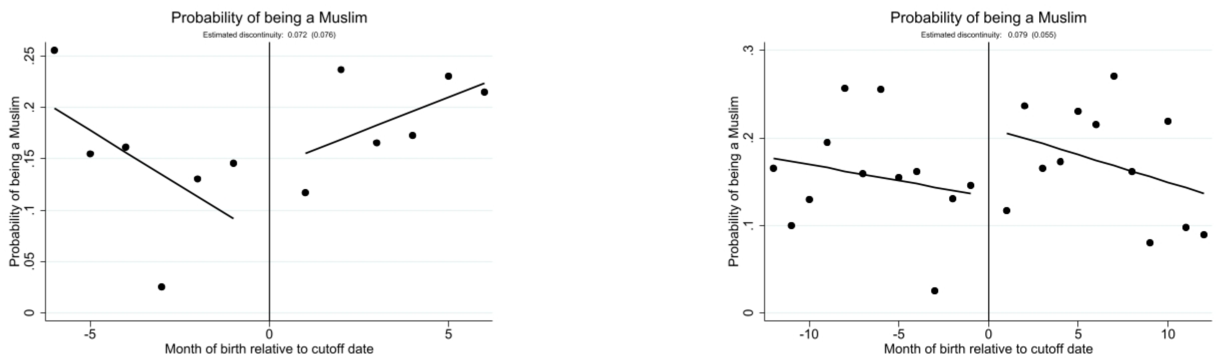


Figure C13: Probability of living in an urban area (Bandwidth Order: 6 months and 12 months)

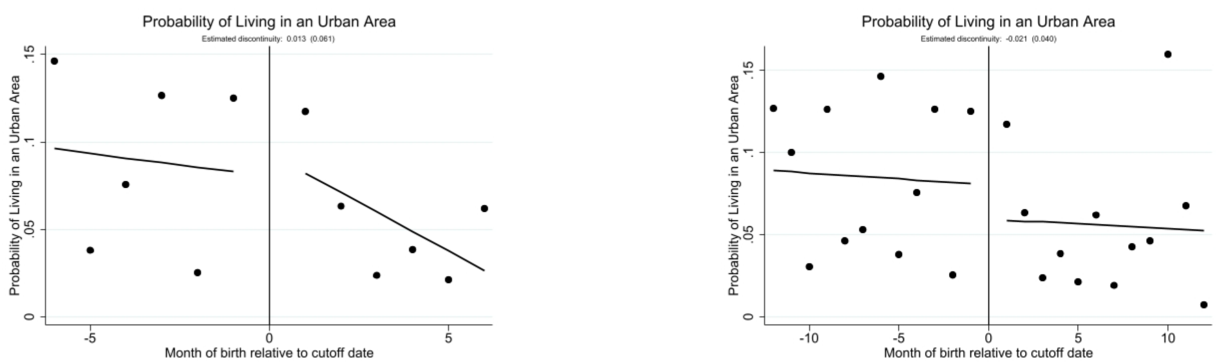
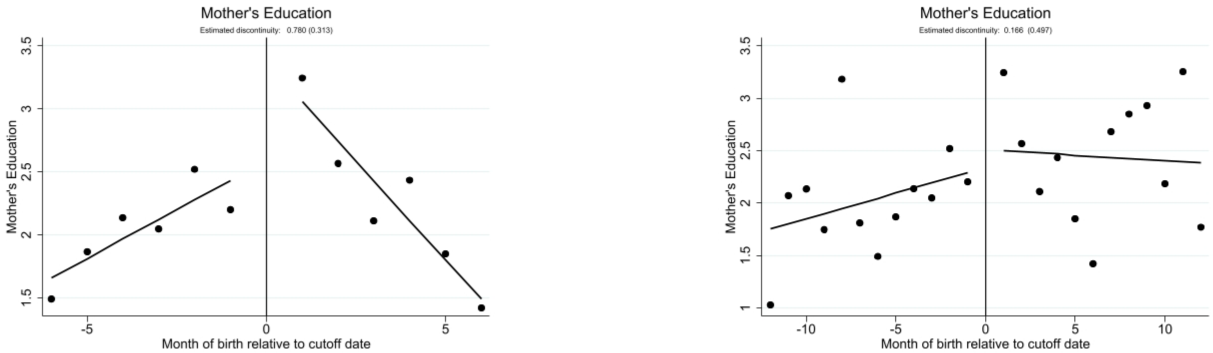
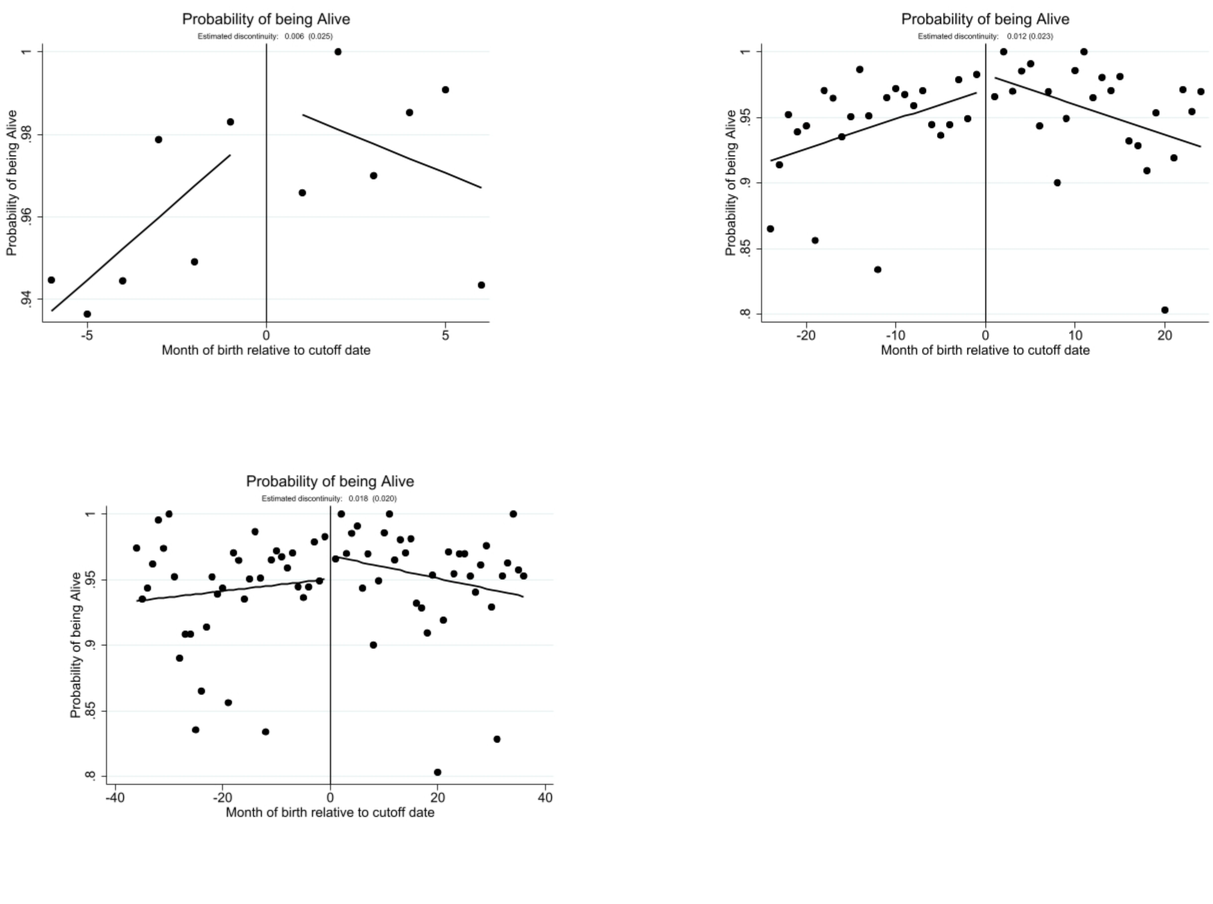


Figure C14: Mother’s Education (Bandwidth Order: 6 months and 12 months)

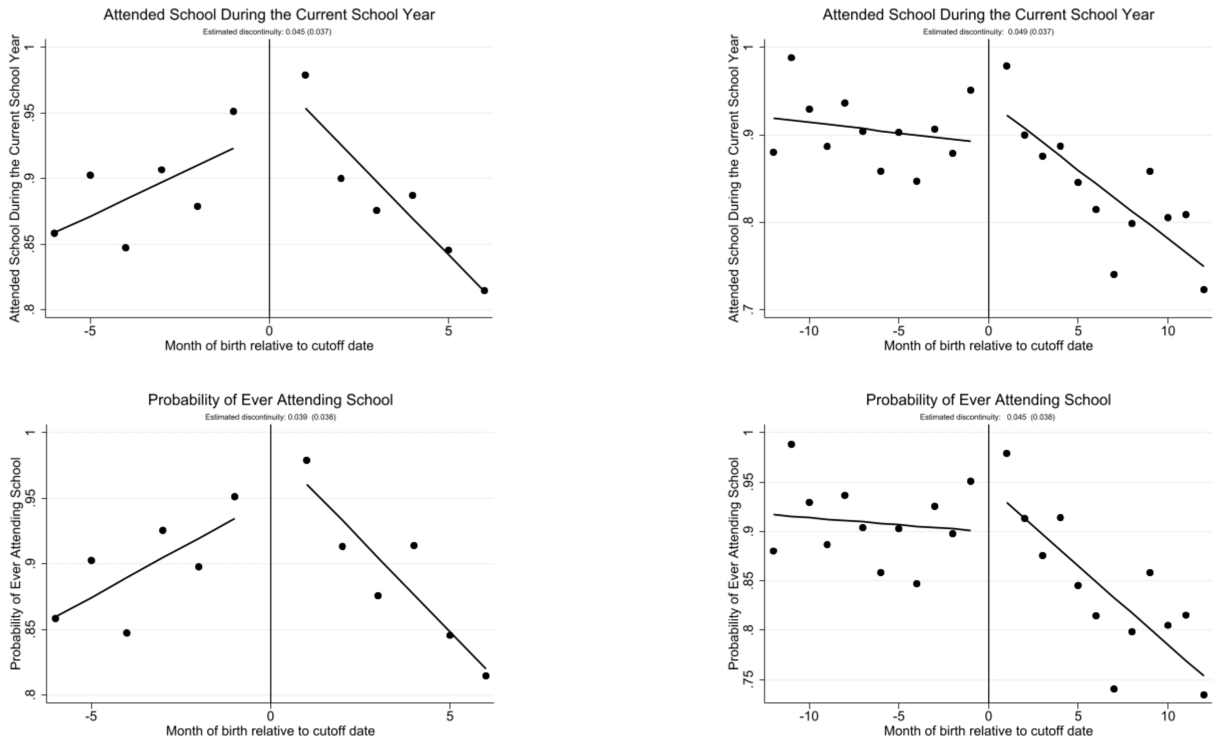


Section 3: Figure C15: Outcome 1- Probability of Survival of the Girl Child [41] (Bandwidth Order: 6 months, 18 months, and 36 months)

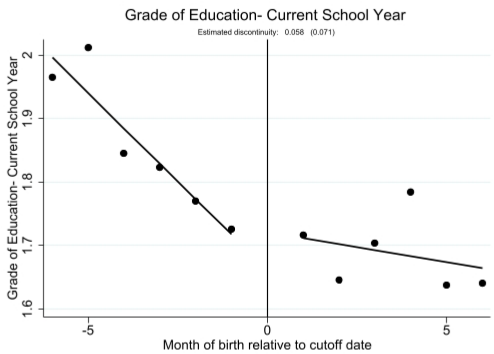


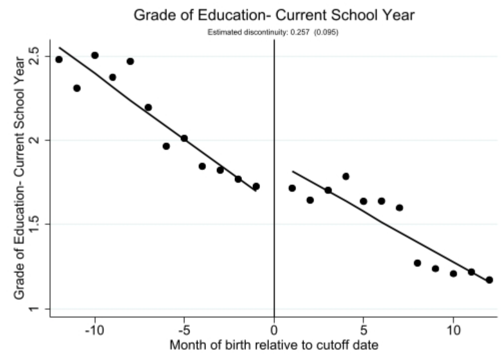
[41] The estimated discontinuity is computed using linear regression ; robust standard errors in parenthesis; the fitted values from this regression are plotted via lines. Note: 95% Confidence Intervals not plotted due to the existence of state sampling weights in the data; Code for graphs has been taken from Clark and Royer (2013)

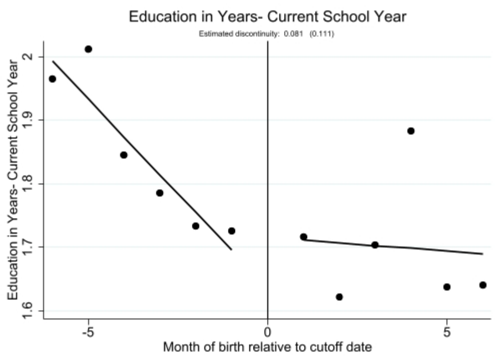
Section 4: Figure C16: Outcome 2- Education Outcomes [42] (Bandwidth Order: 6 months and 12 months)

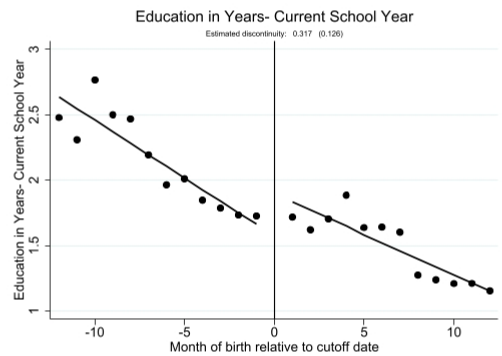


[42] The estimated discontinuity is computed using linear regression ; robust standard errors in parenthesis; the fitted values from this regression are plotted via lines. Note: 95% Confidence Intervals not plotted due to the existence of state sampling weights in the data; Code for graphs has been taken from Clark and Royer (2013)









Section 5: Alternative Specification (flexible linear)

Table C3: Probability the Child is Alive at the Time of the Interview (First Born Child)

	(1)	(2)	(3)
	Alive	Alive	Alive
Estimate	-0.041	0.004	0.008
	(0.036)	(0.023)	(0.018)
Mean	0.969	0.952	0.944
Bandwidth	6 months	18 months	36 months
Observations	363	1106	2141

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C4: Probability the Child is Alive at the Time of the Interview (Second Born Girl When the First Born is a Female)

	(1)	(2)	(3)
	Alive	Alive	Alive
Estimate	0.093	0.063*	0.053**
	(0.087)	(0.034)	(0.024)
Mean	0.964	0.963	0.961
Bandwidth	6 months	18 months	36 months
Observations	156	551	1016

Note: All estimates are weighted using state women's sampling weights; Robust standard errors in parentheses * p < 0.1, **p < 0.05, *** p < 0.01

Table C5: Probability the Child is Alive at the Time of the Interview (Second Born Girl When the First Born is a Male)

	(1)	(2)	(3)
	Alive	Alive	Alive
Estimate	0.015	0.017	0.010
	(0.052)	(0.033)	(0.025)
Mean	0.971	0.970	0.952
Bandwidth	6 months	18 months	36 months
Observations	160	500	953

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C6: Probability the Child Went to School During the Current School Year (First Born Child)

	(1)	(2)
	Attending School	Attending School
Estimate	0.070	0.066
	(0.072)	(0.052)
Mean	0.905	0.881
Bandwidth	6 months	12 months
Observations	338	685

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C7: Probability the Child Ever Went to School (First Born Child)

	(1)	(2)
	Child Ever Went to School	Child Ever Went to School
Estimate	0.081	0.067
	(0.071)	(0.051)
Mean	0.909	0.884
Bandwidth	6 months	12 months
Observations	338	685

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C8: Grade Attended During the Current School Year (First Born Child)

	(1)	(2)
	Grade Attended During the Current School Year	Grade Attended During the Current School Year
Estimate	0.289	0.397*
	(0.298)	(0.215)
Mean	1.747	1.815
Bandwidth	6 months	12 months
Observations	308	606

Note: All estimates are weighted using state women's sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C9: Years of Education Conditional on Ever Attending School (First Born Child)

	(1)	(2)
	Years of Education Conditional on Ever Attending School	Years of Education Conditional on Ever Attending School
Estimate	0.226	0.401*
Mean	1.752	1.831
Bandwidth	6 months	12 months
Observations	310	610

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C10: Probability the Child Went to School During the Current School Year (Second Born Girl When the First Born is a Female)

	(1) Attending School	(2) Attending School
Estimate	-0.073	0.030
	(0.105)	(0.081)
Mean	0.867	0.895
Bandwidth	6 months	12 months
Observations	141	296

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C11: Probability the Child Ever Went to School (Second Born Girl When the First Born is a Female)

	(1) Child Ever Went to School	(2) Child Ever Went to School
Estimate	-0.078	0.025
	(0.105)	(0.081)
Mean	0.869	0.896
Bandwidth	6 months	12 months
Observations	141	296

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C12: Grade Attended During the Current School Year (Second Born Girl When the First Born is a Female)

	(1) Grade Attended During the Current School Year	(2) Grade Attended During the Current School Year
Estimate	-0.475	0.278
	(0.463)	(0.316)
Mean	1.756	1.849
Bandwidth	6 months	12 months
Observations	123	265

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C13: Years of Education Conditional on Ever Attending School (Second Born Girl When the First born is a Female)

	(1) Years of Education Conditional on Ever Attending School	(2) Years of Education Conditional on Ever Attending School
Estimate	-0.466	0.288
	(0.463)	(0.316)
Mean	1.752	1.847
Bandwidth	6 months	12 months
Observations	124	266

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C14: Probability the Child Went to School During the Current School Year (Second Born Girl When the First Born is a Male)

	(1) Attending School	(2) Attending School
Estimate	-0.025	0.003
	(0.128)	(0.088)
Mean	0.881	0.085
Bandwidth	6 months	12 months
Observations	156	344

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C15: Probability the Child Ever Went to School (Second Born Girl When the First Born is a Male)

	(1) Child Ever Went to School	(2) Child Ever Went to School
Estimate	-0.076	-0.006
	(0.121)	(0.083)
Mean	0.900	0.854
Bandwidth	6 months	12 months
Observations	156	344

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C16: Grade Attended During the Current School Year (Second Born Girl When the First Born is a Male)

	(1) Grade Attended During the Current School Year	(2) Grade Attended During the Current School Year
Estimate	-0.069	0.067
	(0.585)	(0.299)
Mean	1.848	1.841
Bandwidth	6 months	12 months
Observations	139	294

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C17: Years of Education Conditional on Ever Attending School (Second Born Girl When the First Born is a Male)

	(1) Years of Education Conditional on Ever Attending School	(2) Years of Education Conditional on Ever Attending School
Estimate	0.037	0.110
	(0.581)	(0.300)
Mean	1.808	1.842
Bandwidth	6 months	12 months
Observations	142	297

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Section 6: Alternative Specification (Flexible Quadratic)

Table C18: Probability the Child is Alive at the Time of the Interview (First Born Child)

	(1) Alive	(2) Alive	(3) Alive
Estimate	-0.041	0.004	0.008
	(0.036)	(0.023)	(0.018)
Mean	0.969	0.952	0.944
Bandwidth	6 months	18 months	36 months
Observations	363	1106	2141

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C19: Probability the Child is Alive at the Time of the Interview (Second Born Girl When the First Born is a Female)

	(1) Alive	(2) Alive	(3) Alive
Estimate	0.095	0.064*	0.052**
	(0.075)	(0.033)	(0.024)
Mean	0.964	0.963	0.961
Bandwidth	6 months	18 months	36 months
Observations	156	551	1016

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C20: Probability the Child is Alive at the Time of the Interview (Second Born Girl When the First Born is a Male)

	(1) Alive	(2) Alive	(3) Alive
Estimate	0.017	0.018	0.011
	(0.050)	(0.033)	(0.025)
Mean	0.971	0.970	0.952
Bandwidth	6 months	18 months	36 months
Observations	160	500	953

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C21: Probability the Child Went to School During the Current School Year (First Born Child)

	(1) Attending School	(2) Attending School
Estimate	0.070	0.061
	(0.072)	(0.052)
Mean	0.905	0.881
Bandwidth	6 months	12 months
Observations	338	685

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C22: Probability the Child Ever Went to School (First Born Child)

	(1) Child Ever Went to School	(2) Child Ever Went to School
Estimate	0.080	0.062
	(0.072)	(0.051)
Mean	0.909	0.884
Bandwidth	6 months	12 months
Observations	338	685

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C23: Grade Attended During the Current School Year (First Born Child)

	(1) Grade Attended During the Current School Year	(2) Grade Attended During the Current School Year
Estimate	0.308	0.310
	(0.297)	(0.205)
Mean	1.747	1.815
Bandwidth	6 months	12 months
Observations	308	606

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C24: Years of Education Conditional on Ever Attending School (First Born Child)

	(1) Years of Education Conditional on Ever Attending School	(2) Years of Education Conditional on Ever Attending School
Estimate	0.289	0.410*
	(0.297)	(0.217)
Mean	1.752	1.831
Bandwidth	6 months	12 months
Observations	310	610

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C25: Probability the Child Went to School During the Current School Year (Second Born Girl When the First Born is a Female)

	(1) Attending School	(2) Attending School
Estimate	-0.009	0.030
	(0.106)	(0.084)
Mean	0.867	0.895
Bandwidth	6 months	12 months
Observations	141	296

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C26: Probability the Child Ever Went to School (Second Born Girl When the First Born is a Female)

	(1) Child Ever Went to School	(2) Child Ever Went to School
Estimate	-0.013	0.025
	(0.106)	(0.084)
Mean	0.869	0.896
Bandwidth	6 months	12 months
Observations	141	296

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C27: Grade Attended During the Current School Year (Second Born Girl When the First Born is a Female)

	(1) Grade Attended During the Current School Year	(2) Grade Attended During the Current School Year
Estimate	-0.681	0.221
	(0.447)	(0.307)
Mean	1.756	1.849
Bandwidth	6 months	12 months
Observations	123	265

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C28: Years of Education Conditional on Ever Attending School (Second Born Girl When the First Born is a Female)

	(1) Years of Education Conditional on Ever Attending School	(2) Years of Education Conditional on Ever Attending School
Estimate	-0.674	0.229
	(0.447)	(0.307)
Mean	1.752	1.847
Bandwidth	6 months	12 months
Observations	124	266

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C29: Probability the Child Went to School During the Current School Year (Second Born Girl When the First Born is a Male)

	(1) Attending School	(2) Attending School
Estimate	-0.012	-0.003
	(0.128)	(0.085)
Mean	0.881	0.085
Bandwidth	6 months	12 months
Observations	156	344

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C30: Probability the Child Ever Went to School (Second Born Girl When the First Born is a Male)

	(1) Child Ever Went to School	(2) Child Ever Went to School
Estimate	-0.061	-0.014
	(0.122)	(0.080)
Mean	0.900	0.854
Bandwidth	6 months	12 months
Observations	156	344

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C31: Grade Attended During the Current School Year (Second Born Girl When the First Born is a Male)

	(1) Grade Attended During the Current School Year	(2) Grade Attended During the Current School Year
Estimate	-0.117	0.068
	(0.568)	(0.299)
Mean	1.848	1.841
Bandwidth	6 months	12 months
Observations	139	294

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C32: Years of Education Conditional on Ever Attending School (Second Born Girl When the First Born is a Male)

	(1) Years of Education Conditional on Ever Attending School	(2) Years of Education Conditional on Ever Attending School
Estimate	-0.021	0.110
	(0.564)	(0.299)
Mean	1.808	1.842
Bandwidth	6 months	12 months
Observations	142	297

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Section 7: Alternative Specification (Local Quadratic)

Table C33: Probability the Child is Alive at the Time of the Interview (First Born Child)

	(1) Alive	(2) Alive	(3) Alive
Estimate	-0.121	-0.040	-0.005
	(0.075)	(0.030)	(0.025)
Mean	0.969	0.952	0.944
Bandwidth	6 months	18 months	36 months
Observations	363	1106	2141

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C34: Probability the Child is Alive at the Time of the Interview (Second Born Girl When the First Born is a Female)

	(1) Alive	(2) Alive	(3) Alive
Estimate	0.211	0.068	0.015
	(0.174)	(0.065)	(0.040)
Mean	0.964	0.963	0.961
Bandwidth	6 months	18 months	36 months
Observations	156	551	1016

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C35: Probability the Child is Alive at the Time of the Interview (Second Born Girl When the First Born is a Male)

	(1)	(2)	(3)
	Alive	Alive	Alive
Estimate	-0.009	-0.014	-0.000
	(0.070)	(0.050)	(0.040)
Mean	0.971	0.970	0.952
Bandwidth	6 months	18 months	36 months
Observations	160	500	953

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C36: Probability the Child Went to School During the Current School Year (First Born Child)

	(1)	(2)
	Attending School	Attending School
Estimate	0.107	0.125
	(0.130)	(0.084)
Mean	0.905	0.881
Bandwidth	6 months	12 months
Observations	338	685

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C37: Probability the Child Ever Went to School (First Born Child)

	(1)	(2)
	Child Ever Went to School	Child Ever Went to School
Estimate	0.116	0.142*
	(0.130)	(0.083)
Mean	0.909	0.884
Bandwidth	6 months	12 months
Observations	338	685

Note: All estimates are weighted using state women's sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C38: Grade Attended During the Current School Year (First Born Child)

	(1)	(2)
	Grade Attended During the Current School Year	Grade Attended During the Current School Year
Estimate	-0.006	0.269
	(0.520)	(0.326)
Mean	1.747	1.815
Bandwidth	6 months	12 months
Observations	308	606

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C39: Years of Education Conditional on Ever Attending School (First Born Child)

	(1)	(2)
	Years of Education Conditional on Ever Attending School	Years of Education Conditional on Ever Attending School
Estimate	-0.174	0.265
	(0.536)	(0.328)
Mean	1.752	1.831
Bandwidth	6 months	12 months
Observations	310	610

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C40: Probability the Child Went to School During the Current School Year (Second Born Girl When the First Born is a Female)

	(1) Attending School	(2) Attending School
Estimate	-0.119	-0.106
	(0.201)	(0.105)
Mean	0.867	0.895
Bandwidth	6 months	12 months
Observations	141	296

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C41: Probability the Child Ever Went to School (Second Born Girl When the First Born is a Female)

	(1) Child Ever Went to School	(2) Child Ever Went to School
Estimate	-0.112	-0.110
	(0.201)	(0.105)
Mean	0.869	0.896
Bandwidth	6 months	12 months
Observations	141	296

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C42: Grade Attended During the Current School Year (Second Born Girl When the First Born is a Female)

	(1) Grade Attended During the Current School Year	(2) Grade Attended During the Current School Year
Estimate	-0.848	-0.050
	(0.850)	(0.498)
Mean	1.756	1.849
Bandwidth	6 months	12 months
Observations	123	265

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C43: Years of Education Conditional on Ever Attending School (Second Born Girl When the First Born is a Female)

	(1) Years of Education Conditional on Ever Attending School	(2) Years of Education Conditional on Ever Attending School
Estimate	-0.862	-0.042
	(0.849)	(0.498)
Mean	1.752	1.847
Bandwidth	6 months	12 months
Observations	124	266

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C44: Probability the Child Went to School During the Current School Year (Second Born Girl When the First Born is a Male)

	(1) Attending School	(2) Attending School
Estimate	-0.064	-0.050
	(0.222)	(0.138)
Mean	0.881	0.085
Bandwidth	6 months	12 months
Observations	156	344

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C45: Probability the Child Ever Went to School (Second Born Girl When the First Born is a Male)

	(1) Child Ever Went to School	(2) Child Ever Went to School
Estimate	-0.061	-0.014
	(0.122)	(0.080)
Mean	0.900	0.854
Bandwidth	6 months	12 months
Observations	156	344

Note: All estimates are weighted using state women's sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C46: Grade Attended During the Current School Year (Second Born Girl When the First Born is a Male)

	(1) Grade Attended During the Current School Year	(2) Grade Attended During the Current School Year
Estimate	0.730	0.329
	(0.982)	(0.551)
Mean	1.848	1.841
Bandwidth	6 months	12 months
Observations	139	294

Note: All estimates are weighted using state women's sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C47: Years of Education Conditional on Ever Attending School (Second Born Girl When the First Born is a Male)

	(1) Years of Education Conditional on Ever Attending School	(2) Years of Education Conditional on Ever Attending School
Estimate	0.945	0.496
	(0.990)	(0.549)
Mean	1.808	1.842
Bandwidth	6 months	12 months
Observations	142	297

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Section 8: Adding Baseline Household Characteristics

Table C48: Probability the Child is Alive at the Time of the Interview (First Born Child)

	(1) Alive	(2) Alive	(3) Alive
Estimate	-0.045	-0.001	0.007
	(0.034)	(0.023)	(0.018)
Mean	0.969	0.952	0.944
Bandwidth	6 months	18 months	36 months
Observations	363	1106	2141

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C49: Probability the Child is Alive at the Time of the Interview (Second Born Girl When the First Born is a Female)

	(1) Alive	(2) Alive	(3) Alive
Estimate	0.096	0.068**	0.053**
	(0.081)	(0.034)	(0.024)
Mean	0.964	0.963	0.961
Bandwidth	6 months	18 months	36 months
Observations	156	551	1016

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C50: Probability the Child is Alive at the Time of the Interview (Second Born Girl When the First Born is a Male)

	(1)	(2)	(3)
	Alive	Alive	Alive
Estimate	0.026	0.016	0.008
	(0.048)	(0.034)	(0.026)
Mean	0.971	0.970	0.952
Bandwidth	6 months	18 months	36 months
Observations	160	500	953

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C51: Probability the Child Went to School During the Current School Year (First Born Child)

	(1)	(2)
	Attending School	Attending School
Estimate	0.045	0.045
	(0.051)	(0.051)
Mean	0.905	0.881
Bandwidth	6 months	12 months
Observations	338	685

Note: All estimates are weighted using state women's sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C52: Probability the Child Ever Went to School (First Born Child)

	(1)	(2)
	Child Ever Went to School	Child Ever Went to School
Estimate	0.061	0.047
	(0.069)	(0.051)
Mean	0.909	0.884
Bandwidth	6 months	12 months
Observations	338	685

Note: All estimates are weighted using state women's sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C53: Grade Attended During the Current School Year (First Born Child)

	(1)	(2)
	Grade Attended During the Current School Year	Grade Attended During the Current School Year
Estimate	0.238	0.290
	(0.271)	(0.195)
Mean	1.747	1.815
Bandwidth	6 months	12 months
Observations	308	606

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C54: Years of Education Conditional on Ever Attending School (First Born Child)

	(1)	(2)
	Years of Education Conditional on Ever Attending School	Years of Education Conditional on Ever Attending School
Estimate	0.226	0.401*
	(0.272)	(0.212)
Mean	1.752	1.831
Bandwidth	6 months	12 months
Observations	310	610

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C55: Probability the Child Went to School During the Current School Year (Second Born Girl When the First Born is a Female)

	(1) Attending School	(2) Attending School
Estimate	0.040	0.065
	(0.096)	(0.083)
Mean	0.867	0.895
Bandwidth	6 months	12 months
Observations	141	296

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C56: Probability the Child Ever Went to School (Second Born Girl When the First Born is a Female)

	(1) Child Ever Went to School	(2) Child Ever Went to School
Estimate	0.036	0.060
	(0.096)	(0.083)
Mean	0.869	0.896
Bandwidth	6 months	12 months
Observations	141	296

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C57: Grade Attended During the Current School Year (Second Born Girl When the First Born is a Female)

	(1) Grade Attended During the Current School Year	(2) Grade Attended During the Current School Year
Estimate	-0.644	0.219
	(0.442)	(0.305)
Mean	1.756	1.849
Bandwidth	6 months	12 months
Observations	123	265

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C58: Years of Education Conditional on Ever Attending School (Second Born Girl When the First Born is a Female)

	(1) Years of Education Conditional on Ever Attending School	(2) Years of Education Conditional on Ever Attending School
Estimate	-0.659	0.201
	(0.440)	(0.313)
Mean	1.756	1.849
Bandwidth	6 months	12 months
Observations	124	266

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C59: Probability the Child Went to School During the Current School Year (Second Born Girl When the First Born is a Male)

	(1) Attending School	(2) Attending School
Estimate	0.016	0.035
	(0.136)	(0.086)
Mean	0.881	0.085
Bandwidth	6 months	12 months
Observations	156	344

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C60: Probability the Child Ever Went to School (Second Born Girl When the First Born is a Male)

	(1)	(2)
	Child Ever Went to School	Child Ever Went to School
Estimate	-0.043	0.024
	(0.129)	(0.080)
Mean	0.900	0.854
Bandwidth	6 months	12 months
Observations	156	344

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C61: Grade Attended During the Current School Year (Second Born Girl When the First Born is a Male)

	(1)	(2)
	Grade Attended During the Current School Year	Grade Attended During the Current School Year
Estimate	-0.192	0.090
	(0.568)	(0.309)
Mean	1.848	1.841
Bandwidth	6 months	12 months
Observations	139	294

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01

Table C62: Years of Education Conditional on Ever Attending School (Second Born Girl When the First Born is a Male)

	(1)	(2)
	Years of Education Conditional on Ever Attending School	Years of Education Conditional on Ever Attending School
Estimate	-0.071	0.142
	(0.566)	(0.310)
Mean	1.808	1.842
Bandwidth	6 months	12 months
Observations	142	297

Note: All estimates are weighted using women's state sampling weights; Robust standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01