Title

Evaluating Conditional Cash Transfers via Synthetic Control: Estimating Education and Health Effects ofColombia’s Familias en Acción

Author
Wright, Mitchell

Publication Date
2019

Undergraduate

Evaluating Conditional Cash Transfers via Synthetic Control: Estimating Education and Health Effects of Colombia's Familias en Acción

Advisor: Gonzalo Vazquez-Bare

March 18th, 2019

Abstract

Conditional cash transfers have rapidly increased in popularity throughout the past two decades, ushering in a new structure of social welfare programs throughout developing countries. In this thesis, the impacts of Colombia's conditional cash transfer, Familias en Acción, are measured to determine the improvement of both the education and health of Colombia's poorest youth. By using the synthetic control method, estimations are provided at the country level, resulting in easily interpretable findings. When compared to the synthetic control unit proposed in this study, Colombia shows a substantial increase in primary school attendance and a noticeable increase in secondary school attendance. The unique health outcomes observed suggest that Colombia experienced a discernable decrease in child mortality, while impacts on immunization rates remain uncertain.

1 Introduction

In this thesis, the synthetic control method is used to create an estimation of the impacts of Colombia's conditional cash transfer program implemented in 2001, Familias en Acción (FA), on education and health at the country level. Most research conducted on conditional cash transfers focuses on impacts at the municipality level; however, through providing estimations at the country level with the synthetic control method, results are more easily comparable across different programs and countries. The identification of effective programs at the country level is essential for institutions that provide international aid, such as the World Bank Group, which currently estimates that 1.9 billion people receive aid through social safety net programs, with an increasing number coming from conditional cash transfers (World Bank, 2015).

Conditional cash transfers (CCTs) are a type of social welfare program that transfer cash to poor households on the condition that these households have made the prespecified investments in the education and health of their children (Ariel Fiszbien and Norbert Schady, 2009). Unlike unconditional cash transfers, CCTs require poor families to invest in the human capital of their children so the next generation has a pathway out of poverty. Conditional cash transfers have been rapidly gaining popularity with the turn of the 21st century, increasing in prevalence from just 2 countries in 1997 up to 64 countries as of 2015 (World Bank, 2015). With the increasing implementation of CCTs, there is a growing necessity to evaluate the impacts of these programs at the macrolevel so that the effects of various programs can be compared between countries.

Colombia is chosen as the country of interest in this thesis for two main reasons. First, relative to the rest of South America, Colombia implemented its conditional cash transfer early on, which alongside Brazil's conditional cash transfer program, served as a model for many CCT programs in neighboring countries. Second, Colombia's early implementation of Familias en Acción also makes it an apt candidate for analysis through the synthetic control method. In the absence of a clear control group, the synthetic control method establishes a valid control based off a weighted average of sufficiently similar countries without a CCT, which is then used to compare against Colombia. Thus, the differences between the synthetic weighted average and Colombia reveal the impact of FA in the years following its implementation.

To identify the impact of FA on the education and health of the impoverished youth in Colombia, the following education and health outcomes are observed: primary gross enrollment, secondary gross enrollment, under-5 mortality, and DPT immunization rates.

By establishing a valid control group through the application of the synthetic control method, it is demonstrated that following the implementation of FA in the year 2001, primary gross enrollment experienced an average increase of roughly 17% in the following 5 years when compared to the synthetic control unit. Effects in secondary gross enrollment are found to be smaller at an average increase of around 2.6%. When observing health outcomes, under-5 mortality rates show a meaningful decrease of nearly 300 fewer deaths per 1,000 live births by the year 2006. Impacts on DPT immunization rates are more ambiguous with a 2.9% average increase associated with the implementation of the program in the following 5 years.

The rest of this thesis is organized as follows: the remainder of Section 1 provides an overview of Colombia's Familias en Acción (FA) as well as a brief review of related research, Section 2 details the synthetic control method and explains randomization inference, Section 3 describes the data being used, Section 4 presents the results for each outcome of interest as well as a discussion pertaining to limitations that may impact these results, and finally, a brief conclusion is stated in Section 5. An appendix is included at the end containing tables for all synthetic control weights as well as relevant robustness checks.

1.1 Familias en Acción

In 2001, the Colombian government implemented a conditional cash transfer program, known as Familias en Acción, that reached 700 out of the 1,102 municipalities in Colombia. In 2006, the program included nearly 500,000 households and had a budget of roughly US$140 million (Francisco Ayala, 2006). Created as a response to the Colombian economic crisis during the late 1990's, FA aims to "help reduce, overcome, and prevent poverty and income inequality, accumulate human capital, and improve living conditions of poor and vulnerable families through complementing their income" (Nadin Medellín and Fernando Sánchez Prada, 2015). The specific target population of the program is families with children ages 0-17 and a household income in the lowest quintile (bottom 20%) of the national income distribution as identified by the Beneficiary Identification System for Social Programs (SISBEN).

Cash transfers ranging in size from US$4 to US$25 are made once every two months to eligible families on the condition that the families meet the prespecified education and health requirements detailed by the program. Educational goals are focused on increasing school enrollment. The specific education requirements state that children must be enrolled in either primary school (ages 7-11) or secondary school (ages 11-17) and attend 80% of school days at minimum. Health goals are more broadly centered around general improvements in overall health, including vaccinations/immunizations. The health requirements state that children (ages 0-7) must attend regular growth and development check-ups at a local healthcare provider (Medellín and Prada, 2015).

1.2 Related Research

Most evaluations of conditional cash transfers, including those done on Familias en Acción, are conducted at the microlevel, typically by comparing treated municipalities (those affected by the CCT) to control municipalities (those unaffected by the CCT) within a specific country. Attanasio, Battistin, Fitzsimons, Mesnard, and Vera-Hernández (2005) follow this general procedure by using a standard difference-in-differences model to evaluate the impact of FA in Colombia. They estimate that FA had a small, non-statistically significant increase in primary net attendance of around 1%, and a substantial statistically significant increase in secondary net attendance of around 7% to 8%. They attribute the insignificant findings in primary net attendance to the fact that prior to the implementation of FA, primary school attendance was already close to 100%.

However, while Attanasio, et al. (2005) use net attendance rates, this study observes gross enrollment rates, which allow for enrollment percentages to rise above 100% due the incorporation of individuals outside of their official school-age grade. This is an important addition, and its relevance can be interpreted as follows: children who join the school system late (for example, individuals who may have been absent from the school system until the implementation of FA) may be placed in a lower grade than their official school-age would dictate due to their late academic start, and it is important to use a metric that accounts for these individuals when education impacts are being observed. Nonetheless, most studies, such as García and Hill (2009) as well as Zavakou (2011), find that FA has generally positive improvements in education outcomes in treated municipalities.

Unlike education outcomes, health outcomes are evaluated with less frequency and typically provide inconclusive results at best. In Attanassio, et al. (2005), the frequency of health visits for children under 12 months of age shows a slight increase; however, no indicators of actual health improvements show conclusive results. This study observes unique health outcomes in addition to education outcomes to provide a comprehensive analysis of the program's impact in Colombia.

Approaching the evaluation of FA with a different technique, Cigliutti, Echeverri Gómez, Golinksy, Gutiérrez, and Sorá (2015) implement the synthetic control method to estimate the impact of FA on secondary gross enrollment in Colombia, in which an average of a 5% increase is attributed to the program. However, Cigliutti, et al. (2015) provide a broad surface level analysis of multiple South American programs in which many important characteristics of Familias en Accíon are not discussed. In this study, a more scrupulous implementation of the synthetic control method specific to Colombia's Familias en Acción is utilized to provide an in-depth estimation of the impacts of the program.

The methods implemented in this study harken back to Abadie, Diamond, and Hainmueller (2010), in which the synthetic control method is applied to estimate the impacts of a tobacco tax increase in California. The importance of synthetic control is that it allows for the use of largescale aggregate data to estimate the impacts of a policy on a region as a whole. The macrolevel estimation of FA's impact in both education and health presented in this study is intended to complement existing literature by providing a comprehensive analysis of the program that is easily compared between different countries and their corresponding conditional cash transfer programs.

2 Methodology

This section provides a brief explanation of the synthetic control method, its implementation in this study, and the process used to interpret the results. For a more comprehensive explanation of the synthetic control method, refer to Abadie and Gardeazabal (2003) or Abadie, Diamond, Hainmuller (2010).

2.1 Model: Synthetic Control

Synthetic control is often used when trying to identify the impact of a large-scale treatment (program, policy, intervention, etc.) on a specific outcome in a unit (region, state, country, etc.) in which the treatment is believed to be the predominant influencer on the outcome in observation within the unit. For example, Abadie, et al. (2010) implement synthetic control to observe the impact of California's tobacco tax passed in 1988 on per capita cigarette sales in California in the 10 years following the tax. This method proves especially advantageous when aggregate data is being used in place of a large quantity of micro data to observe effects at a macrolevel. The underlying principal motivating synthetic control is that by comparing outcomes in similar units that did not receive treatment (control units) to the outcome in the single unit that did receive treatment (treated unit), conclusions can be made regarding the impact of the treatment on the observed outcome.

Selecting Control Units

Often times any one control unit, no matter how similar it may be to the treated unit, cannot accurately represent the counterfactual of what would have occurred if the treated unit never received treatment. Synthetic control bypasses this problem by constructing a synthetic unit created from a weighted average of a collection of control units (also referred to as the donor pool) that depicts an accurate counterfactual representation of the treated unit had it never received treatment. In terms of Abadie, et al. (2010), a synthetic California is created using a donor pool composed of states that have similar per capita cigarette sales to California, but did not experience a tobacco tax increase. Thus, the weighted average of the donor pool is able to represent what would have occurred in California if the tobacco tax was not increased.

To create an accurate synthetic unit, numerous safeguards must be adhered to when selecting the control units to be used in the donor pool. First, it is essential that no control units in the donor pool have a similar treatment to the one being implemented in the treated unit. This is because the weighted average of the donor pool is supposed to represent what would have happened in the treated unit in the absence of the treatment. If any of the control units in the donor pool have an equivalent treatment, then the resulting synthetic unit will be biased and unable to accurately depict a counterfactual of the treated unit. In the case of the tobacco tax study, states with similar tobacco tax increases are disqualified from the donor pool because they cannot help construct a synthetic California that represents California in the absence of an increased tobacco tax.

A second but equally important qualification for the control units in the donor pool is that they must match the treated unit in outcome trends. To quantify the impacts of the treatment, it is necessary to select relevant outcomes to observe before and after the treatment so that the impact of the treatment can be identified. Again, in terms of the tobacco tax study, Abadie, et al. (2010) use per capita cigarette sales as the relevant outcome to observe because any effect that the tax increase might have on cigarette consumption should be reflected by per capita cigarette sales.

The general trends in the chosen outcome must be similar between the treated unit and the control units so that the weighted average of the donor pool represents the treated unit as closely as possible. If the outcomes in observation have vastly different patterns between the treated unit and control units (for example, increasing rates for the treated unit versus decreasing rates for the control units over time), then no combination of control units will be able to accurately depict the trends of the outcome in the treated unit. This plays an important role in the creation of the weighted average, since not all control units in the donor pool are assigned equal weight.

Constructing the Weighted Average

Once the appropriate donor pool has been created, each of the control units in the donor pool is given a specific weight depending on the similarity in outcome trends to the treated unit. Control units with outcome trends that are more closely related to the treated unit are given more weight so that the most accurate synthetic unit is constructed.

To validate that certain control units in the donor pool should receive higher weighting than others, a set of covariates (referred to as predictors) are analyzed as well. These predictors are characteristics of the treated unit that may also be impacting the outcome in observation. Valid predictors may include any factors that are related to the outcome but are not directly influenced by the implementation of the treatment. For example, in the tobacco tax study, beer consumption per capita, percent of the population ages 15-24, and a few other relevant predictors are selected due to their close relation to the outcome being observed, per capita cigarette sales, while simultaneously remaining unaffected by any direct implications of the tobacco tax increase.

Predictors are important because a synthetic unit constructed based on the outcome of interest alone will never provide an absolute perfect match to the treated unit; however, by cross-validating the trends in the outcome of interest with trends in the appropriate predictors, these errors can be mitigated, resulting in a more accurate depiction of the counterfactual treated unit by the synthetic unit (Robert Bifulco, Ross Rubenstein, and Hosung Sohn, 2017).

Pre/Post Treatment Periods and Mean Squared Prediction Error (MSPE)

With these prior conditions satisfied, the synthetic unit created from the weighted average of the donor pool should provide an accurate representation of the treated unit. The mean squared prediction error (MSPE) provides insight into both the accuracy of the synthetic model as well as the impact of the treatment. The MSPE is the difference between treated unit and the synthetic unit and is observed in 2 different periods; prior to the treatment and after the treatment. Before the treatment is implemented, the trends of the treated unit and the synthetic unit should be as similar as possible, resulting in a low pretreatment MSPE. After the treatment is implemented, assuming that the treatment has a noticeable effect on the outcome in observation, the treated unit and the synthetic unit will begin to diverge. In this posttreatment period, a large difference between the treated unit and synthetic unit suggests a large impact associated with the treatment, denoted by a high posttreatment MSPE. Since the synthetic unit should closely resemble the treated unit during the pretreatment period, the difference in outcomes in the posttreatment period can be attributed as the impact of the treatment.

2.2 Implementation

In this thesis, the synthetic control method is used to assess the impact of Colombia’s conditional cash transfer, Familias en Acción, on the education and health of the youth targeted by this program. Familias en Acción is implemented in 2001, so all years prior to 2001 are considered the pretreatment period, and all years after and including 2001 are considered the posttreatment period. In this study, the posttreatment period extends to 2006. It is within this 5-year window that the effects of FA are observed.

Below, the observed outcomes of interest, the countries in the donor pool, and the predictors used to create the synthetic control model are listed. For specific definitions of outcomes and predictors, refer to Section 3, Data.

- **Outcomes** (4)
  - Education
    - Primary gross enrollment
    - Secondary gross enrollment
  - Health
    - Under-5 mortality
    - DPT Immunization
- **Donor pool*** (13)
  - South America
    - Guyana
    - Venezuela
  - Central America
    - Belize
    - Guatemala
    - Panama
  - Global matches
    - Albania
    - Algeria
    - Botswana
    - Fiji
    - Libya
    - Namibia
    - Romania
    - South Africa
- **Predictors** (13)
  - General
    - GDP per capita
    - GNI per capita
    - Gross Savings
    - Urban population
    - Rural population
    - Population ages 0-14
  - Education
    - Expenditure on primary education
    - Expenditure on secondary education
  - Health
    - Health expenditure per capita
    - Access to electricity
    - Access to at least basic drinking water services
    - Access to at least basic sanitation services
    - Life expectancy at birth

*Global matches are selected based off similar historical income classification to Colombia. None of the countries in the donor pool have a CCT or a similar program to Famlias en Acción in the time period being observed. In addition, these countries are also verified as having similar trends to Colombia with respect to one or more outcomes of interest. This is displayed graphically in Section 3.1, Preliminary Visualization.

2.3 Interpretation: Randomization Inference

Unlike typical regression-based models, the accuracy of the synthetic control model is not derived from computing standard errors and running large sample inferential techniques. To verify the credibility of the proposed counterfactual presented by the synthetic unit, a type of permutation test is applied so that randomization inference can be performed.

The permutation test, or placebo test, takes each of the countries in the donor pool and runs iterations of the model one by one with each unit designated as the treated unit. Recall that none of the countries in the donor pool have a similar CCT program (i.e., received "treatment"), but that they do share many similarities to Colombia in regards to the outcomes of interest. Because none of the countries in the donor pool have a similar program implemented, the posttreatment MSPE should suggest little measurable impact in any of the placebo tests. If the model were to be dictated by randomness or noise, then it is just as likely for any one of the countries in the donor pool to show posttreatment effects comparable to those of Colombia since the measured effect would not be resulting from Familias en Acción, but rather from random fluctuations in trends.

The randomization inference process is illustrated through two distinct graphs presented in Section 4. The first graph is a placebo plot that shows the differences between each placebo test and its respective synthetic control as well as Colombia and synthetic Colombia. The placebo plot displays the results graphed over time, distinguishing between the pretreatment and posttreatment periods.

The second graph quantifies the difference between the observed impact in Colombia and the placebo tests by plotting the ratios of post MSPE to pre MSPE (posttreatment impact relative to pretreatment error) for each country. A high MSPE ratio is achieved by a high posttreatment effect and a low pretreatment error. Thus, if the findings of the model are significant, Colombia should have one of the highest MSPE ratios and be substantially differentiated from the average.

3 Data

To estimate the impact of Familias en Acción on education and health, annual time series data for Colombia and all the countries in the donor pool are first gathered from the World Bank data catalog. The observed outcomes of interest for education and health come from the two following databanks respectively: Education Statistics and Health, Nutrition, and Population Statistics. Education data range from 1975 to 2006 while health data range from 1992 to 2006 due to data availability.

For education, primary gross enrollment rate and secondary gross enrollment rate are the observed outcomes of interest. Gross enrollment rate depicts the total enrollment in either primary or secondary school, regardless of age, as a percentage of the eligible official school-age population corresponding to primary/secondary education for each grade. These rates can rise above 100% since children outside of their official school-age grade level are accounted for. It is important to account for total enrollment regardless of age because of the assumption that many students enrolling in school as a result of FA may likely be placed in a lower grade than suggested by their official school-age level due to prior absenteeism from the school system.

For health, under-5 mortality rate and DPT immunization rate are the outcomes of interest used to observe the impact of FA. Under-5 mortality rate is the number of children projected to die before reaching the age of 5 per 1,000 live births. DPT immunization rate is the percentage of children ages 12-23 months who have received the vaccination for Diphtheria/Pertussis/Tetanus. DPT vaccination is a standard vaccine administered in developing countries, which occurs in Colombia around 18 months of age (World Health Organization, 2013).

Predictor values are equally essential to the construction of the synthetic control. All predictors are also extracted from the World Bank data catalog, namely, the World Development Indicators databank. There are 13 predictors used to aid the estimation of the synthetic control model. The time period to which the data pertain ranges from 1971 to 2006. The predictors used are the following: GDP per capita (current US $), GNI per capita (current US $), Gross Savings (current US $), urban population (% of total pop.), rural population (% of total pop.), population ages 0-14 (% of total pop.), expenditure on primary education (% of government expenditure on education), expenditure on secondary education (% of government expenditure on education), current health expenditure per capita (current US $), access to electricity (% of total pop.), access to at least basic drinking water services (% of total pop.), access to at least basic sanitation services (% of total pop.), and life expectancy at birth (years).

The first six predictors are general country statistics, the next two are specific to education statistics, and the final 5 are specific to health statistics. This is so that general development trends as well as education/health trends that may be influencing the outcomes of interest are accounted for in the synthetic control model.

3.1 Preliminary Visualization

Below, Figures 1 through 4 show each outcome of interest (primary gross enrollment, secondary gross enrollment, under-5 mortality rate, and DPT immunization respectively) for Colombia as well as for the unweighted average of the donor pool. The unweighted average is calculated by taking a simple arithmetic mean of the donor pool for each outcome. These graphs serve two functions: one is to validate the selection of control units in the donor pool by verifying similarity of trends, and the other is to highlight the necessity of implementing synthetic control to create a better fitting counterfactual. The time period for education outcomes ranges from 1975 to 2006, while the time period for health outcomes ranges from 1992 to 2006. The vertical black line dividing the graphs into pre and post 2001 periods marks the implementation of Familias en Acción.

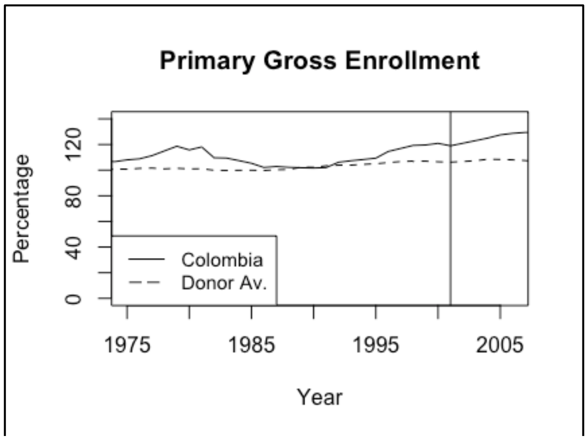

Figure 1. Primary gross enrollment: Colombia vs. unweighted donor average

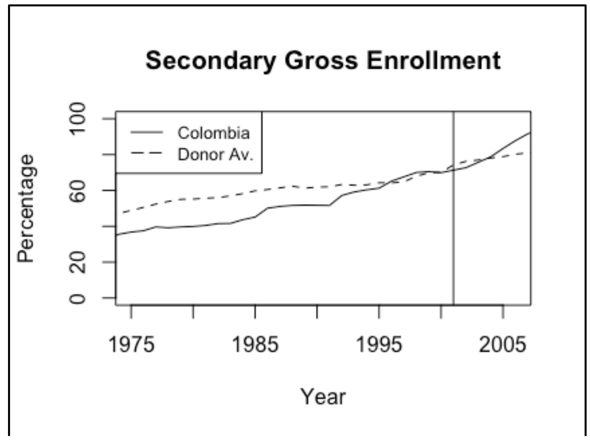

Figure 2. Secondary gross enrollment: Colombia vs. unweighted donor average

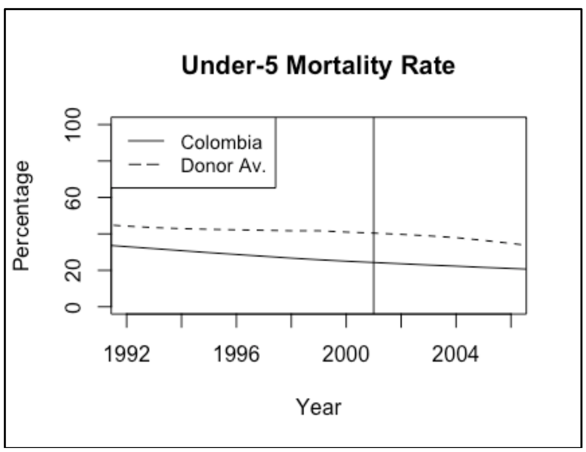

Figure 3. Under-5 mortality: Colombia vs. unweighted donor average

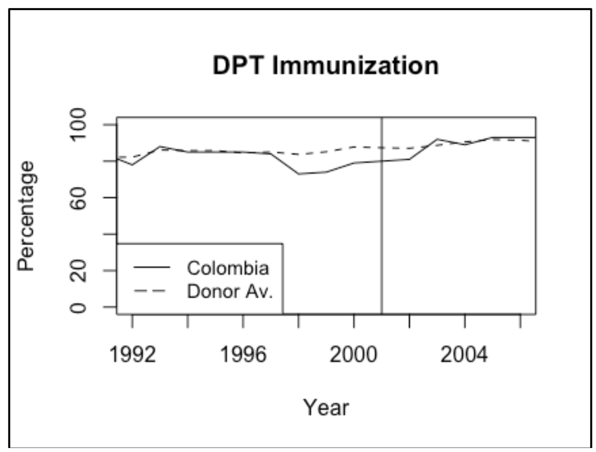

Figure 4. DPT immunization: Colombia vs. unweighted donor average

The general trends between Colombia and the unweighted donor average for each outcome are similar, yet far from identical. The similarity in trends confirms that the control units chosen for the donor pool are indeed comparable to Colombia with respect to each outcome of interest; however, it is not possible to extract any conclusive results from these preliminary graphs since there are substantial differences between Colombia and the unweighted average of the donor pool prior to the implementation of FA in 2001. Therefore, these preliminary graphs emphasize the necessity of implementing synthetic control to create a weighted average to accurately represent a counterfactual Colombia in the absence of FA so that meaningful inference can be conducted. The synthetic control graphs are shown below in Section 4, Results.

4 Results

The results for each outcome are presented in the following manner: first, a graph of the synthetic control model is displayed, second, the placebo plots are shown, and last, a graph of the MSPE ratios is shown. For the first outcome discussed, a detailed explanation of each graph is provided. At the end of this section, certain limitations are discussed.

4.1 Primary Gross Enrollment

Synthetic Control Model

Below, Figure 5 depicts primary gross enrollment rates for both Colombia and synthetic Colombia. The solid black line represents the observed primary gross enrollment rates in Colombia. The dashed black line represents synthetic Colombia, i.e., the primary gross enrollment rates estimated by the weighted average of the donor pool. On the x-axis, time is displayed in years, ranging from 1975 to 2006. The y-axis denotes the percentage of students enrolled in primary school regardless of official school-age grade. The red line marks the implementation of Familias en Acción in the year 2001, dividing the graph into pretreatment and posttreatment periods.

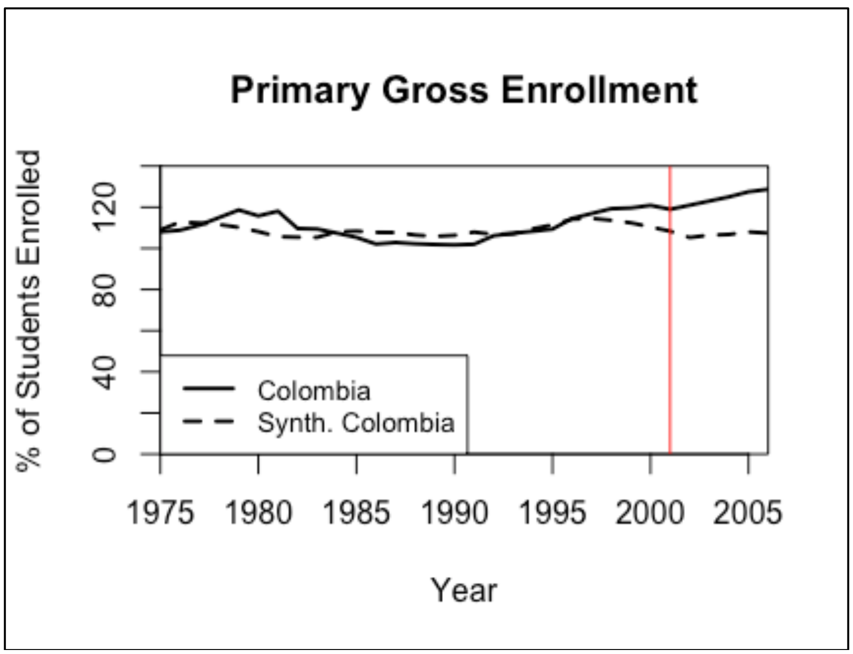

Figure 5. Trends in primary gross enrollment: Colombia vs. synthetic Colombia

Notice that when compared to the unweighted average of the donor pool depicted in Figure 1 from Section 3.1, synthetic Colombia provides a much better pretreatment fit. This fact, combined with the cross-validation of predictors, suggests that synthetic Colombia provides a sensible approximation of the primary gross enrollment rate in Colombia post 2001 in the absence of Familias en Acción.

The pretreatment MSPE is only 5.2% on average, meaning that prior to the implementation of FA, primary gross enrollment rates depicted by synthetic Colombia approximate the actual rates in Colombia with an average error of roughly 5%. This pretreatment MSPE is substantially smaller than the posttreatment MSPE, which is 17% on average, signifying that FA is associated with an increase in primary gross enrollment of roughly 17% in the 5 years following its implementation. Assuming that the 5.2% error extends into the posttreatment period, there is still a substantial increase in primary gross enrollment associated with the implementation of the program.

The largest measured effect is observed in 2006 at 21.3%, which suggests that the impacts of FA are growing over time. This makes sense logically since it takes a few years for a large-scale program like Familias en Acción to develop and reach its entire target audience.

An important detail to address is the slight divergence of synthetic Colombia from Colombia a few years prior to 2001, starting around 1995. This difference is largely explained by the education reform policies enacted in Colombia around this time, namely, the Decentralization Bill in 1994. As a result of this bill, municipalities were granted a larger budget to allocate towards the social sector, resulting in an increase in spending on education (Pamela Lowden, 2011). However, this was a short-term increase, and as such, it is unlikely that it accounts for a substantial amount of the observed impact in later years following the implementation of Familias en Acción.

The estimated impact of Familias en Acción on primary gross enrollment rates is substantially large when compared to previous findings in primary *net* enrollment rates due to the nature of the metric. Gross enrollment rates account for individuals outside of their school-age specified grade, whereas net enrollment rates do not. A large increase in primary gross enrollment associated with FA might suggest that many children who enroll as a result of the program are entering into lower grades than their school-age would dictate since they are behind their peers in years of formal education.

Randomization Inference

Below, Figure 6 shows the gaps between primary gross enrollment of Colombia and synthetic Colombia from the previous graph (Figure 5) as well as the gaps for each individual placebo test. The x-axis displays the same time period as Figure 5 (1975 to 2006). The y-axis is the magnitude of the gap in percentage. The dashed line marks the implementation of FA in 2001, again, dividing the graph into pretreatment and posttreatment periods.

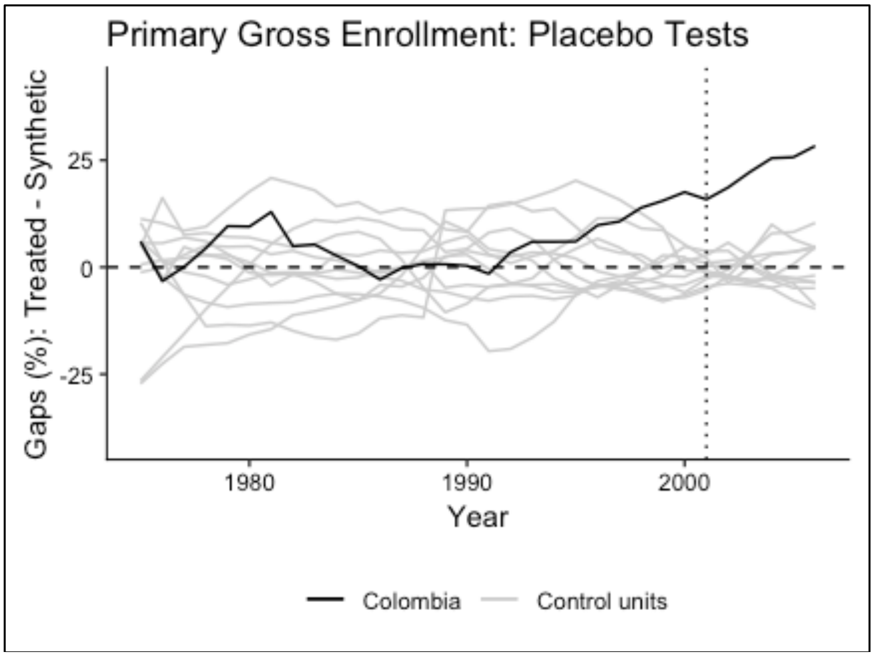

Figure 6. Gaps in primary gross enrollment: Colombia and control units

This graph shows the effects measured by the synthetic model for Colombia as well as each control country. The grey lines ("Control units") represent each of the placebo plots that result from applying the synthetic control model to each of the countries in the donor pool. That is, the grey lines show the difference in primary gross enrollment rates between each country in the donor pool and its respective synthetic control.

Since Colombia is the only country that actually implemented a conditional cash transfer in this time period, it should be the only unit that shows a substantial impact in the posttreatment period. This is clearly confirmed by the graph, as Colombia shows a much larger treatment effect than any of the placebo tests after the implementation of FA. If the impact being measured by the model had no correlation with the implementation of a CCT, then it would be just as likely that the placebo plots would show posttreatment trends similar to that of Colombia since the effect is not dictated by the treatment. This is evidently not the case, as the placebo plots oscillate around zero in the posttreatment period, suggesting minimal effect and confirming that the impact measured by the synthetic model is attributed to FA and is not just a result of outside noise.

Below, Figure 7 provides the primary gross enrollment post MSPE to pre MSPE ratios, denoting the posttreatment effect relative to the pretreatment error for synthetic Colombia and each respective synthetic control. These ratios are plotted along the x-axis with their respective country labeled on the y-axis.

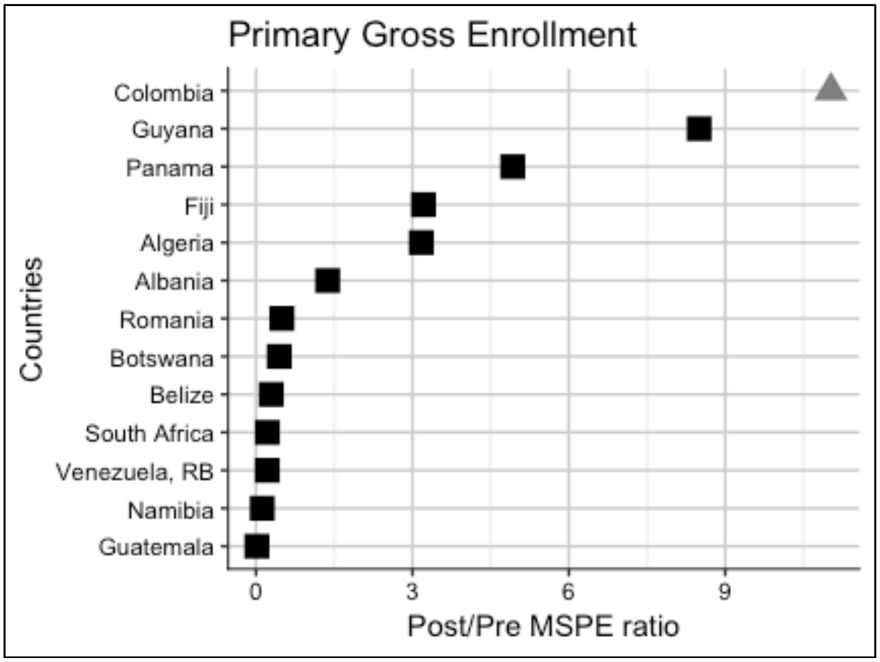

Figure 7. Ratio of posttreatment MSPE to pretreatment MSPE: Colombia and control units

Colombia's Post/Pre MSPE-ratio stands out from the rest by a noticeable amount. The post MSPE is about 11 times the pre MSPE for synthetic Colombia, whereas for the other synthetic controls, the average ratio is roughly 1.9. This reinforces the fact that the measured treatment effects are resulting primarily from Familias en Acción, and not by some outside noise. If other Post/Pre MSPE-ratios were of equivalent size to Colombia on average, then the measured effects could not be confidently attributed to the implementation of FA. However, this is not the case, as Colombia is substantially differentiated from the mean.

An important note to address is that there are only 12 countries out of the 13 in the donor pool used in Figure 7. For primary gross enrollment specifically, Libya has been excluded in the construction of the synthetic unit due to its dissimilarity of trends. The selection of countries from the donor pool varies between outcomes in order to utilize the most appropriate countries for the specific outcome in observation. The reason Libya is not discarded from the donor pool altogether is that it is essential in the creation of the synthetic unit for the health outcomes observed in this study. The tables detailing the specific weight assigned to the control units in the donor pool used to create the synthetic unit for each outcome is available in the appendix.

4.2 Secondary Gross Enrollment

Synthetic Control Model

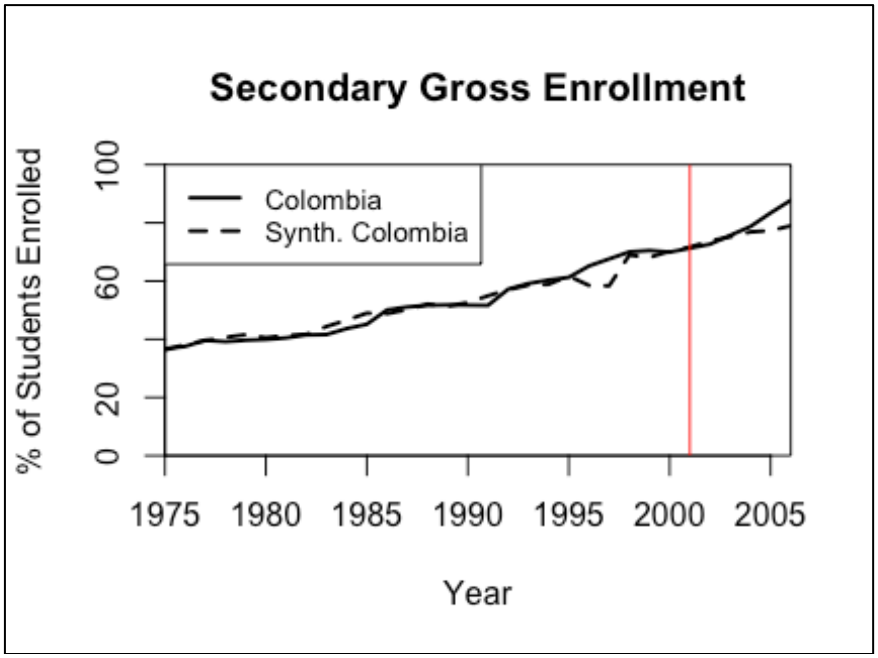

Figure 8. Trends in secondary gross enrollment: Colombia vs. synthetic Colombia

In the pretreatment period, synthetic Colombia provides an accurate depiction of secondary gross enrollment rates in Colombia. The closeness in fit is a result of the smoothness in trends observed in secondary gross enrollment rates in Colombia; there is a constant upward trend with minimal fluctuation. This makes it easier for the synthetic unit to emulate since there are no idiosyncrasies specific to Colombia that the weighted average must try to match.

On average, there is a very low pretreatment MSPE of 1.7%. However, the posttreatment MSPE is only 2.6%, suggesting that there is minimal impact on secondary gross enrollment associated with FA in the 5 years after its implementation. The largest observed effect is measured in 2006, at 8.8%, which similar to primary gross enrollment, can be attributed to the fact that the program may take time to operate at its potential. Nonetheless, Familias en Acción seems to have a substantially smaller impact on gross enrollment rates in secondary school than it does in primary school.

This is presumably due to two related reasons. Unlike primary school, secondary school is divided into upper and lower schooling, of which only lower secondary is compulsory and provided for free. Secondary gross enrollment rates do not distinguish between upper and lower secondary, and since upper secondary school has a direct enrollment cost, it's possible that low-income families are dissuaded from paying the fee and decide to remove their child from school.

Furthermore, individuals in upper secondary school are at the age where they would likely acquire a job and earn an income in the absence of schooling, this poses an indirect opportunity cost to the family by forgoing an additional source of income. When joined together, these adverse factors regarding the cost of secondary school can outweigh the incentive provided by Familias en Acción, resulting in fewer families enrolling their child in secondary school.

Randomization Inference

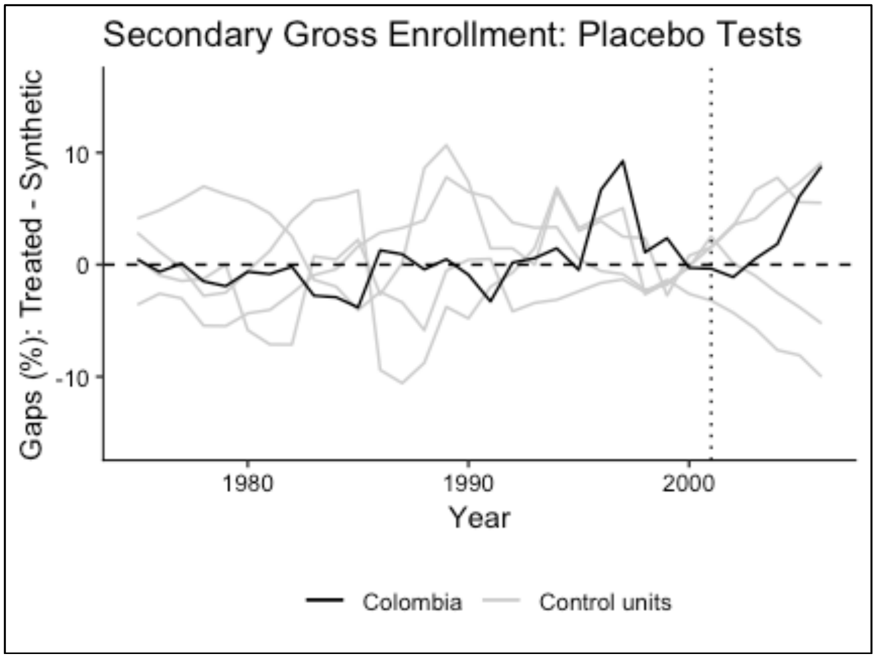

Figure 9. Gaps in secondary gross enrollment: Colombia and control units

There are noticeably fewer placebo plots present in Figure 9 than in Figure 6. This is because countries whose synthetic unit have an unreasonably high pre MSPE do not provide useful information to measure the rarity of observing a large post MSPE and are excluded from the graph. Countries with a pre MSPE greater than 5 times that of Colombia are chosen to be excluded since past this cutoff, the synthetic model no longer provides an accurate representation of the pretreatment period for these countries and therefore they should not be used to make inferences about the posttreatment period.

As is evident in the graph, the posttreatment effect observed in Colombia is not substantially differentiated from those of the placebo tests. As stated before, the impact on secondary gross enrollment associated with FA is not very large to begin with, therefore there is an increased likelihood that control countries can show effects of a similar size purely based on randomness of trends in secondary gross enrollment. These values are quantified and discussed in detail with Figure 10.

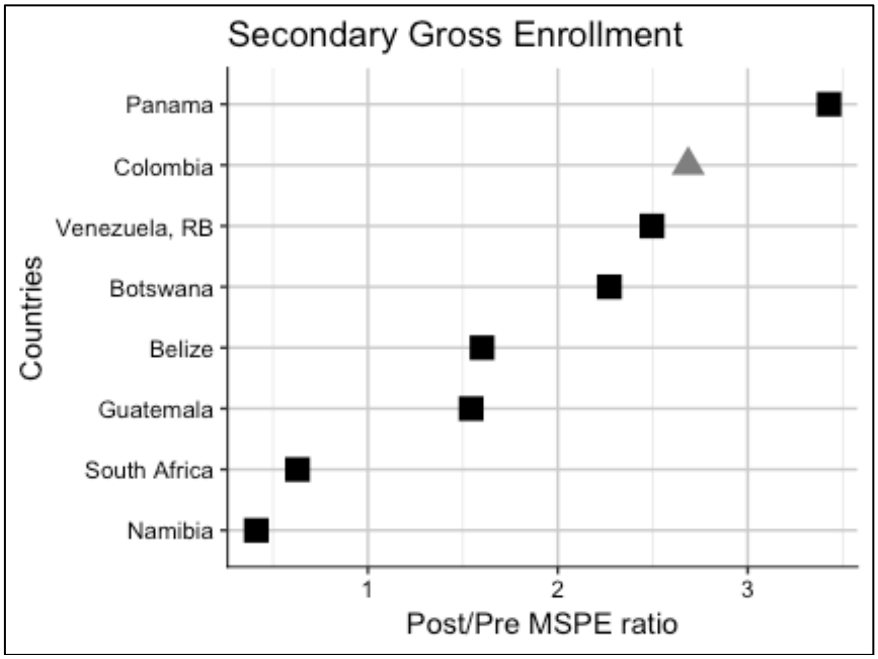

Figure 10. Ratio of posttreatment MSPE to pretreatment MSPE: Colombia and control units

In Figure 10, Colombia's Post/Pre MSPE-ratio is not the largest, nor is it substantially differentiated from the average. Colombia's post MSPE is about 2.7 times larger than its pre MSPE, while the average ratio is around 1.8. Since other countries are showing almost the same impact relative to their error, the results presented by the model for Colombia are not as significant. As a result, it is harder to confidently attribute the majority of the observed increase in secondary gross enrollment rates to Familias en Acción since control countries are experiencing similar increases without an equivalent program.

4.3 Under-5 Mortality

Synthetic Control Model

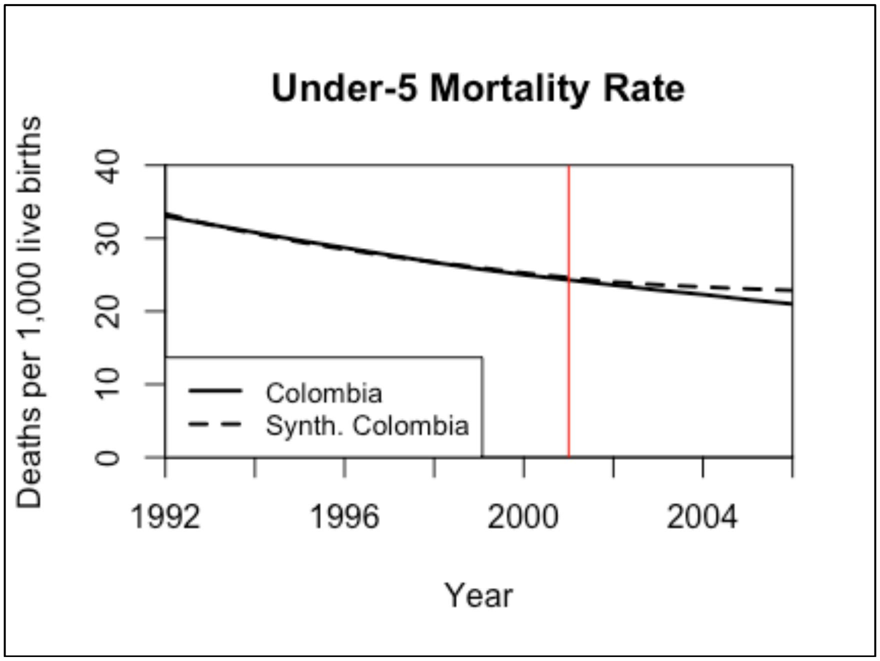

Figure 11. Trends in under-5 mortality: Colombia vs. synthetic Colombia

For under-5 mortality, synthetic Colombia displays an even more accurate representation of the observed rates in Colombia. In terms of mortalities per 1,000 live births, the average pretreatment MSPE is .18 compared to a posttreatment MSPE of .97. Unlike the education outcomes, the posttreatment effect observed for under-5 mortality is expressed as a decrease, signifying that per 1,000 live births, there are .97 fewer under-5 mortalities associated with the implementation of FA in the following 5 years. The largest decrease is observed in 2006, at 1.86 fewer under-5 mortalities per 1,000 live births.

This impact might seem negligible at first glance, however, once put into perspective, the resulting effect is pronounced. In 2006, there were 18.7 live births per 1,000 people in Colombia, resulting in roughly 160,000 lives births in the 1st quintile (World Data Atlas, 2006). As a result, this impact equates to nearly 300 fewer under-5 mortalities per year.

Randomization Inference

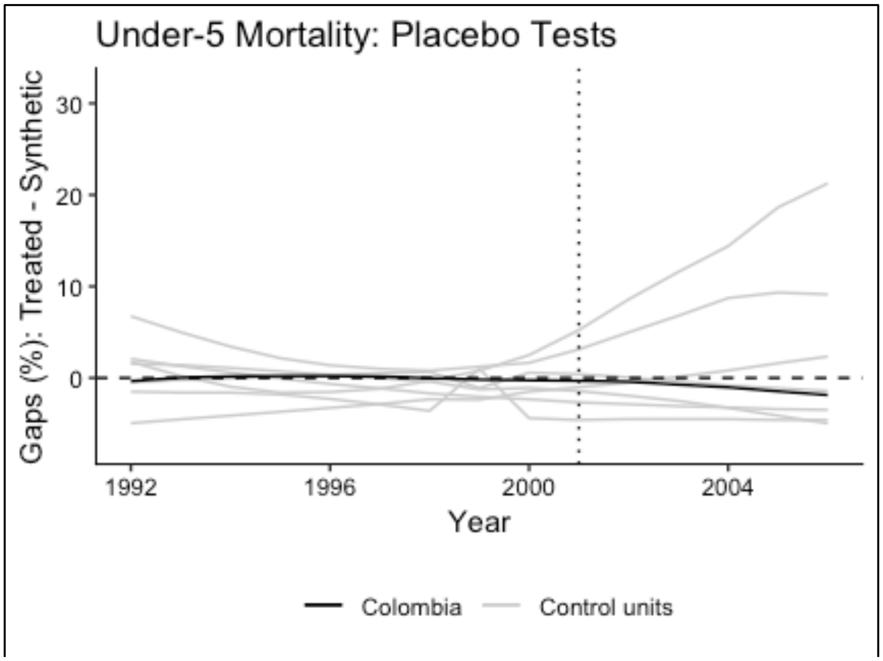

Figure 12. Gaps in under-5 mortality: Colombia and control units

In Figure 12, three control countries show a greater posttreatment impact than Colombia, however, none of them have a pretreatment error that is lower than Colombia. Thus, the placebo plot needs to be examined with the Post/Pre MSPE-ratios in order to make any formal conclusions regarding the significance of the measured impact. This placebo plot does however serve one essential function; highlighting outliers. While most control countries see a moderate decrease in under-5 mortality in the posttreatment period, one control country experiences a sharp increase in under-5 mortality rates. Since none of the control countries have a similar program implemented in the period in observation, this is due to random fluctuations specific to the individual country and does not bias the effects estimated for Colombia. The impact of the outlying control country is discussed in conjunction with Figure 13, which distinguishes it more clearly.

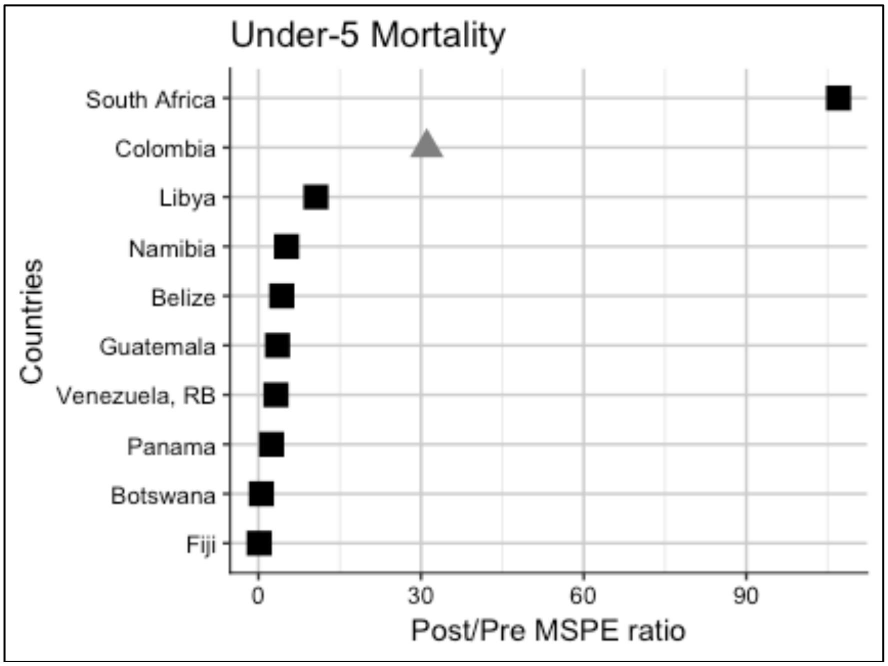

Figure 13. Ratio of posttreatment MSPE to pretreatment MSPE: Colombia and control units

South Africa is a clear outlier in Figure 13, with a Post/Pre MSPE-ratio nearly 3 times greater than that of Colombia. However, the ratios displayed only account for the magnitude of the measured impact, not the directionality. This is why it is important to include the placebo plots as a cross-reference to determine if any outliers are detrimental to the significance of the model. By referencing Figure 12, it is seen that South Africa experiences an *increase* in under-5 mortalities unassociated with any effects of Familias en Acción, and as such, it's high Post/Pre MSPE-ratio does not affect the significance of the model.

As a general robustness check, a test is run excluding South Africa from the synthetic control model for under-5 mortality rates. The pretreatment MSPE increases nominally, but the results remain pronounced. South Africa is ultimately left in the donor pool because upon cross-referencing it with Figure 12, it does not pose any issues to the interpretation of the model.

With this in mind, Colombia has the largest valid Post/Pre MSPE-ratio by a substantial amount. Colombia's post MSPE is about 31 times larger than its pre MSPE. The average ratio of the remaining countries is roughly 3.4. This leads to the conclusion that the measured effects are a result of Familias en Acción and not just noise.

4.4 DPT Immunization

Synthetic Control Model

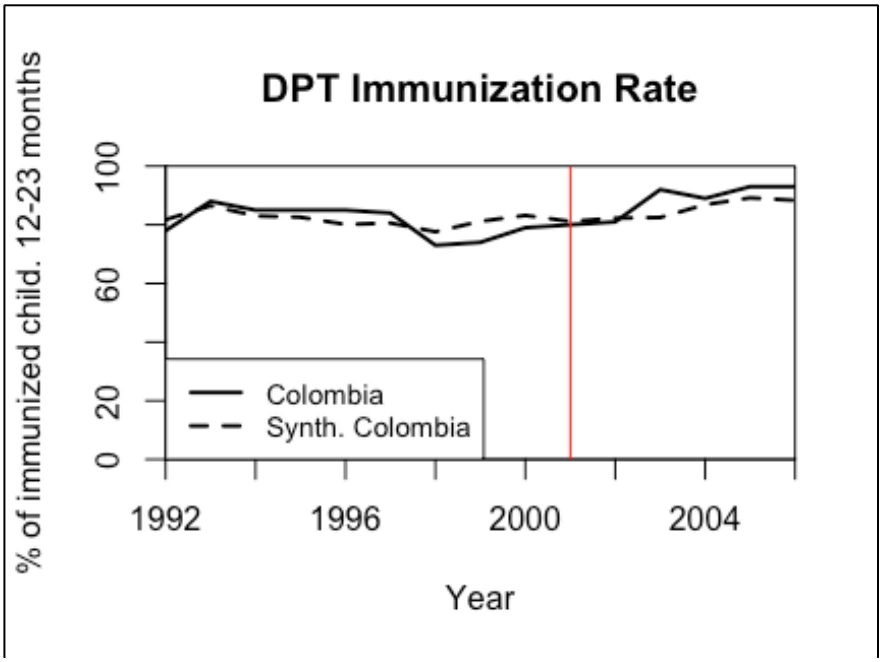

Figure 14. Trends in DPT Immunization: Colombia vs. synthetic Colombia

The synthetic model provides a reasonable pretreatment fit of DPT immunization rates in Colombia, with an average pretreatment MSPE of 3.8%. The posttreatment impact, however, fluctuates greatly and has an average posttreatment MSPE of only 2.9%. In the posttreatment period, the synthetic unit starts off with a higher rate, only to be quickly surpassed in 2003. The differences between synthetic Colombia and Colombia then begin to diminish as time goes on, obscuring any consistent treatment effects.

The ambiguity of Familias en Acción's impact on DPT immunization displayed by the synthetic model is a result of simultaneous influence from exogenous factors. Namely, the creation of Vaccination Week in the Americas, beginning in 2003 in Colombia. This large-scale program enacted by the Pan American Health Organization focuses on administering vaccinations to poor communities as well as raising general awareness of proper vaccination (Pan American Health Organization, 2014). Since different regions of Colombia and each country in the donor pool are each uniquely affected by this program at different times and on different scales, a substantial amount of noise is interjected into the synthetic control model. As a result, it is difficult to confidently attribute any of the measured impact solely to FA.

Randomization Inference

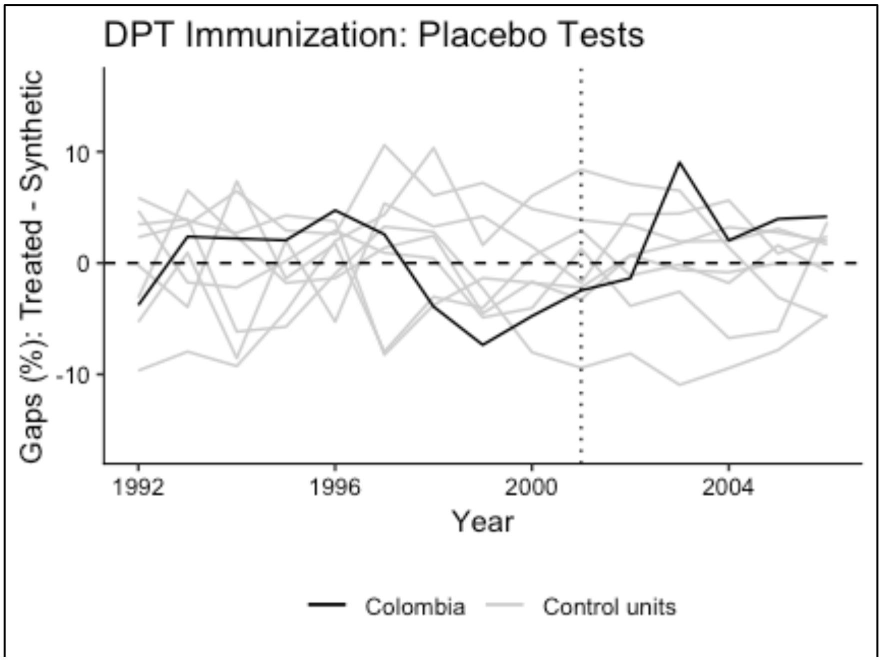

Figure 15. Gaps in DPT Immunization: Colombia and control units

In Figure 15, the placebo tests show a consistent fit during the pretreatment period, confirming that the synthetic model is achieving an accurate approximation during this period. However, the posttreatment effect observed in Colombia is not substantially different than those observed in the control countries. Since numerous countries in the donor pool are also affected by Vaccination Week in the Americas, they are exhibiting a relative increase in DPT immunization rates of equal size to Colombia. This confounds the measured effect and makes it difficult to distinguish how much of the impact is resulting from Familias en Acción and how much is simply spill-over from the increased prevalence of vaccinations efforts stemming from Vaccination Week in the Americas.

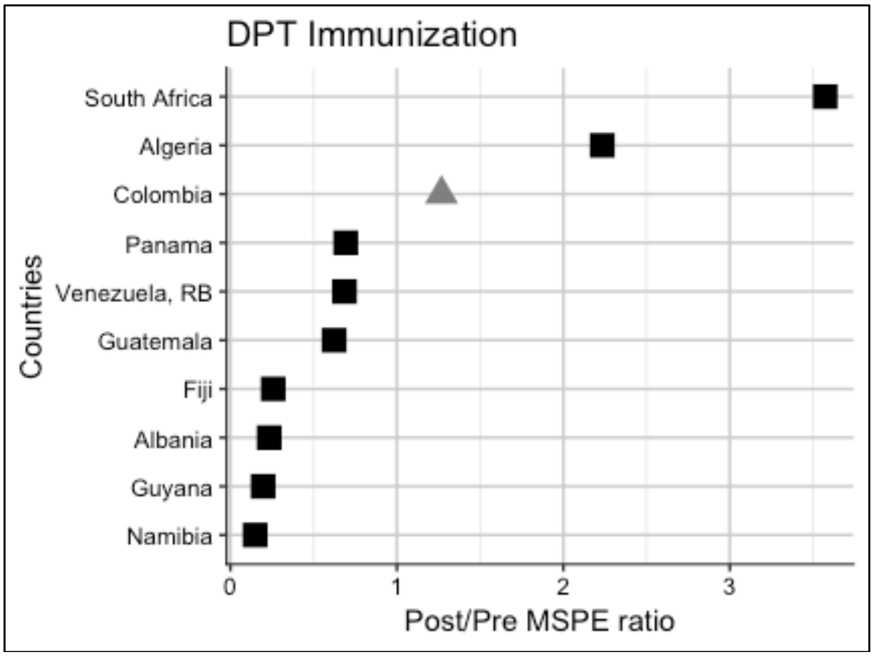

Figure 16. Ratio of posttreatment MSPE to pretreatment MSPE: Colombia and control units

As made apparent in Figure 16, Colombia's Post/Pre MSPE-ratio is not rare in terms of magnitude. Colombia's posttreatment MSPE is about 1.7 times larger than its pretreatment MSPE, but the average ratio is roughly 1, suggesting that the measured effect is not a result unique to FA's impact in Colombia.

4.5 Discussion/Limitations

Since the intention of the synthetic control model is to detect the impact of Familias en Acción at the country level, the outcomes in observation reflect the rates for the population of Colombia as a whole, not just the proportion of the population targeted by the program. As such, there is an expected amount of outside noise entering into the model through families that remain unaffected by the program. However, these confounding effects are mitigated because there are no identifiable programs that would influence the observed outcomes for Colombians unaffected by FA. As a result, the trends of the remaining 80% of the population remain relatively unchanged after the implementation of FA, and any nominal changes that might occur are equally likely to occur in the countries in the donor pool. In future studies, provided the necessary data availability, a similar model could be constructed using data for only the 1st quintile of Colombia and the control countries to diminish any extra noise derived from the remaining population.

Another present limitation is the amount of countries eligible for the donor pool. An abundant donor pool has the potential to allow for a more accurate synthetic unit as well as more placebo tests to verify the significance of the findings. Ideally, more South American countries could be used in the donor pool since these countries are natural candidates to compare to Colombia. However, many South American countries began implementing their own CCTs in the years following 2001, rendering them unqualified for the donor pool. Central American countries are the next suitable choice and are included along with valid global matches to further increase the size of the donor pool. However, it should be noted that the quality of the donor pool cannot be determined purely by size; each country selected for the donor pool must be verified as an appropriate control unit in order to create an accurate synthetic control. There is an inherent tradeoff between the amount of control units in the donor pool and the desired posttreatment window. As the posttreatment window is extended, many suitable control units are disqualified from the donor pool since countries begin implementing their own CCTs throughout the 21st century. In this thesis, a 5-year posttreatment period is suitable because it allows for any lagged effects of Familias en Acción to manifest while still allowing for an ample amount of appropriate countries in the donor pool.

5 Conclusion

Through selecting unique education and health outcomes, a comprehensive assessment of the impacts of Colombia's conditional cash transfer program are presented. By using the synthetic control method, these estimations are done at the country level and allow for an easy comparison of program impacts to be made across countries.

In the 5 years following the implementation of Familias en Acción, the estimated impact on primary gross enrollment is roughly 17%. These substantial findings are largely attributed to an increase in enrollment of children in grades outside of their official school-age level. For secondary gross enrollment, less conclusive results are found. FA is associated with an average increase of 2.6% in the 5 years after its implementation. The small increase in secondary gross enrollment relative to that of primary gross enrollment is presumably due to both the direct and indirect enrollment costs of secondary school. For health outcomes, Familias en Acción is associated with roughly 300 fewer under-5 mortalities per year while impacts on DPT immunization remain uncertain.

These estimations suggest overall positive results from Colombia's conditional cash transfer program and allow for future programs to be compared against Familias en Acción. Conditional cash transfers continue to be a primary tool in the fight against cyclical poverty across the globe, and as such, it is essential that the macrolevel impacts of these programs can be effectively measured.

REFERENCES

Abadie, A., Diamond, A., Hainmueller, J. (2010), "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," *American Statistical Association,* Vol. 105, pp. 493-505.

Abadie, A., and Gardeazabal, J. (2003), "The Economic Costs of Conflict: A Case Study of the Basque Country," *American Economic Review,* pp. 112-32.

Attanasio, O., Battistin, E., Fitzsimons, E., Mesnard, A., and Vera-Hernández, M. (2005), "How Effective are Conditional Cash Transfers? Evidence from Colombia," Institute for Fiscal Studies, pp. 1-11.

Ayala, F. (2005), "Inter-Regional Inequality Facility: Sharing Policies Across Africa, Asia, and Latin America," Policy Brief, Overseas Development Institute, pp. 1-4

Bifulco, R., Ross, R., Sohn, H. (2017), "Using Synthetic Controls to Evaluate the Effect of Unique Interventions: The Case of Say Yes to Education," *SAGE Publications,* Vol. 41, pp. 593-619

Cigliutti, M., Echeverri Gómezm, M., Golinsky, F., Gutiérrez, A., Sorá, G. (2015), "Conditional Cash Transfer Programs: Measuring the Impact on Education," Working Paper, Universidad Torcuato Di Tella, pp. 1-31

Fiszbien, A., and Schady, N (2009), "Conditional Cash Transfers: Reducing Present and Future Poverty," Washington, DC: World Bank, pp. 1-383.

García, S., Hill, J. (2009), "The Impact of Conditional Cash Transfers on Children's School Achievement: Evidence from Colombia," Working Paper, Universidad de los Andes, Facultad de Economía–Cede, pp. 1-48.

Lowden, P. (2011), "Education Reform in Colombia: The Elusive Quest for Effectiveness," Wilson Center, pp. 1-26.

Medellín, N., Prada, F. (2015), "How Does Más Familias en Accíon Work? Best Practices in the Implementation of Conditional Cash Transfer Programs in Latin America and the Caribbean," Inter-American Development Bank, pp. 1-50.

Pan American Health Organization (2014), "Vaccination: Vaccination Week in the Americas," World Health Organization, pp. 1-13.

World Bank (2015), "The State of Social Safety Nets 2015," Washington, DC: World Bank, pp. 1-185.

World Data Atlas. < https://knoema.com/atlas/Colombia > cited 2 March 2019

World Health Organization (2013), "Immunization Summary: A Statistical Reference Containing Data through 2011," World Health Organization, pp. 1-226.

Zavakou, Alkistis. "An Evaluation of the CCT Program Familias en Acción on Grade Failure and School Drop-out Rates," Ph.D. dissertation, European University Institute, 2011. pp. 1-70.

Appendix A: Weighted Averages

Table 1. Country weights for synthetic Colombia: Primary gross enrollment

Country	Weight	Country	Weight
Albania	.371	Guyana	.014
Algeria	0	Namibia	0
Belize	.501	Panama	.114
Botswana	0	Romania	0
Fiji	0	South Africa	0
Guatemala	0	Venezuela	0

Note: Libya is excluded from synthetic Colombia for primary gross enrollment

Table 2. Country weights for synthetic Colombia: Secondary gross enrollment

Country	Weight	Country	Weight
Belize	.56	Panama	.44
Botswana	0	South Africa	0
Guatemala	0	Venezuela	0
Namibia	0		

Note: Albania, Algeria, Fiji, Guyana, Libya, and Romania are excluded from synthetic Colombia for secondary gross enrollment

Table 3. Country weights for synthetic Colombia: Under-5 mortality

Country	Weight	Country	Weight
Belize	.488	Namibia	0
Botswana	0	Panama	.119
Fiji	.33	South Africa	0
Guatemala	0	Venezuela	.005
Libya	0		

Note: Albania, Algeria, Guyana, and Romania are excluded from synthetic Colombia for under-5 mortality

Table 4. Country weights for synthetic Colombia: DPT immunization

Country	Weight	Country	Weight
Albania	0	Guyana	.266
Algeria	.299	Namibia	0
Fiji	.127	Panama	0
Guatemala	0	South Africa	.125

Note: Belize, Botswana, Libya, Romania, and Venezuela are excluded from synthetic Colombia for DPT immunization

Appendix B: Robustness Checks

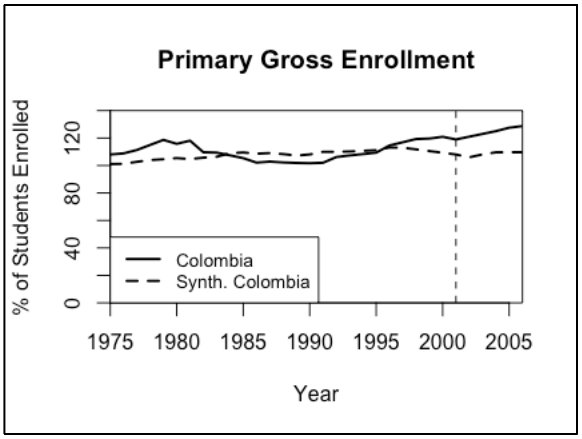

Figure A1. Treatment year assigned to 2006

In Figure A1, the treatment year is assigned to 2006, meaning that the entire time period is the pretreatment period over which the MSPE is minimized. This is to verify that 2001 (the year FA is implemented) is not an arbitrary year in which primary gross enrollment rates just happened to substantially increase. As is evident, the effective increase from 2001 to 2006 remains robust to any changes of the treatment year.

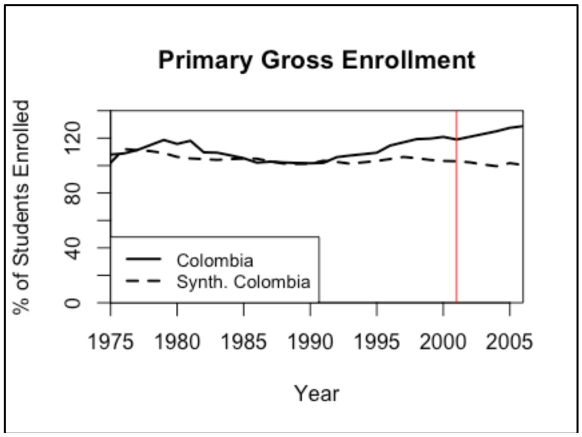

Figure A2. General predictors only

In Figure A2, the synthetic model is shown for primary gross enrollment for which only the 6 general predictors are used. This graph validates the effectiveness of the chosen education and health predictors, and highlights the improved accuracy the results from their inclusion.