UC Irvine
UC Irvine Electronic Theses and Dissertations

Title
Empirical studies on policy evaluation

Permalink
https://escholarship.org/uc/item/7xh9576w

Author
Dona, Gonzalo

Publication Date
2020

Copyright Information
This work is made available under the terms of a Creative Commons Attribution License, availalbe at https://creativecommons.org/licenses/by/4.0/

Peer reviewed|Thesis/dissertation

eScholarship.org
Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA,
IRVINE

Empirical studies on policy evaluation

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Economics

by

Gonzalo Dona

Dissertation Committee:
Professor Matthew Freedman, Chair
Professor Daniel Bogart
Associate Professor Yingying Dong

2020

© 2020 Gonzalo Dona

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank my wife, mother and grandparents for all they have done for me. I would like to thank professor Matthew Freedman, Dan Bogart, and Yingying Dong for their invaluable help with this dissertation. I appreciate the economic assistance from University of California, Irvine and the Chilean government through the scholarship Becas Chile.

VITA

Gonzalo Dona

EDUCATION

Doctor of Philosophy in Economics **2020**
University of California, Irvine *Irvine, California*

Master of Arts in Economics **2017**
University of California, Irvine *Irvine, California*

Master of Arts in Economics **2012**
Universidad Alberto Hurtado *Santiago, Chile*

Bachelor of Arts in Business Administration **2008**
Universidad de los Andes *Santiago, Chile*

RESEARCH EXPERIENCE

Research Assistant **2013–2015**
Universidad Mayor *Santiago, Chile*

TEACHING EXPERIENCE

Teaching Assistant **2016–2020**
University of California, Irvine *Irvine, California*

Lecturer **2014–2015**
Universidad Mayor *Santiago, Chile*

REFEREED CONFERENCE PUBLICATIONS

Mothers' Labor Supply and Conditional Cash Transfer: Evidence from Chile **April 2019**
PAA

ABSTRACT OF THE DISSERTATION

Empirical studies on policy evaluation

By

Gonzalo Dona

Doctor of Philosophy in Economics

University of California, Irvine, 2020

Professor Matthew Freedman, Chair

This dissertation contains three papers that increase our understanding of the impacts of public policies, with a special focus on policies geared toward the poor. The first chapter considers the impact of the minimum wage, one of the most pervasive policy in existence today, to a population that has received very little direct attention in relation to said policy, the unemployed. In this chapter, we show that the policy does not benefit the unemployed. The second chapter studies a welfare program that has seen it's popularity surge in the last couple decades: conditional cash transfers (CCT) are a type of public policy that benefit the poor with cash in exchange for investments in their children. In this chapter, I show that a conditional cash transfer program can have pernicious labor supply distortions if maintained too long. The last chapter focuses not on one policy but rather on the knowledge necessary to create good country-level policies in a increasingly interconnected world. I show in this paper that the key determinant of long term development for a country is its people's principles (i.e., basic ideas), which means that imposing institutions to a third country (i.e., democracy) would be ineffective in increasing their wellbeing.

In the first chapter, we use four panels of the Survey of Income and Program Participation (SIPP) to find how the minimum wage impact unemployed workers, using a difference-in-difference strategy. We take advantage of the high-frequency of the SIPP data to study

both the impact of minimum wage levels at the beginning of an unemployment period and the possible impact of a minimum wage increase while the worker is looking for a job. Our analysis shows that minimum wage levels have relatively mild effects on workers' labor outcomes; it does not show an impact on unemployment duration, or re-employment wages and hours over the next two years. However, minimum wage increases seem to have a more important impact on unemployed workers, their unemployment periods become longer and their working hours' trajectories are worsened. We find it particularly important that our analysis reveals no significant positive effects of higher minimum wages, or its increases, on the unemployed.

The second chapter of this dissertation uses data from seven editions of Chile's National Survey of Socioeconomic Characterization (CASEN), a longitudinal survey meant to evaluate public policies and describe the poor, to show how the oldest conditional cash transfer (CCT) in existence impacts labor force participation of women. Using a difference-in-difference strategy, I evaluate whether a small cash transfer that is suddenly made increasingly more generous and popular after 2007 impacts labor choices by Chilean mothers. Because I study a program that is sixteen years older than any other CCT, my findings can be relevant to rethink newer CCTs; the education level of the population I study more closely resembles educational attainment for many current CCT beneficiaries than their own country's average educational attainment in the early 2000s. I find evidence that older women increase slightly their labor force participation, but younger mothers between 18 to 24 years of age decrease their probability of working significantly (about 4%). The latter finding raises the question of whether the policy should be addressed to this group of young mothers at all.

In the last chapter of this dissertation, I consider the key determinants of long term wellbeing. Using data borrowed from Acemoglu et al. (2002) and complemented with additional relevant information on former colonies I compiled, I show that the aforementioned authors oversimplified the colonization process and that my approach explains the current and historical data better. My theory complicates matters by considering both direction and speed of progress. The direction of progress I argue is determined by principles, which are 'basic' ideas (i.e., a principle: 'all men are created equal'; not a principle: choice of jurisprudence). I show how Europeans played an important role in transmitting these principles (or not) to their colonies; I proxy this transmission with data on founding of the first university in each former colony. My theory also allows for different speeds in progressing towards societal goals. Borrowing from Besley and Persson (2009), I contend that progress is faster for less heterogeneous societies everything else equal, I use population density in 1500 to find the former colonies that would be homogeneous today . My results show that having European principles explains a higher GDP by PPP in 1995 by 0.9 standard deviations, and being homogeneous explains an increase by 1.1 standard deviations in the same metric.

Chapter 1

Effects of the minimum wage on unemployment duration and re-employment outcomes

The employment effects of the minimum wage policy is a very contentious issue that has amassed a very large body of literature. However, almost none of this effort has gone into evaluating how the policy affects unemployed workers directly. Nevertheless, the unemployed are workers that are both in particular need of attention just by being so, and generally over represent less productive workers, who tend to take part in less stable jobs than their more productive counterparts. Additionally, if higher minimum wages have any significant impact on the labor markets, its the unemployed workers that are the most likely to benefit or suffer from it.

For our analysis we use a sample of unemployment spells built from the Survey of Income and Program Participation (SIPP) to establish the relationship between minimum wages and various outcomes of unemployed individuals. More precisely, we study the relationship minimum wages have with the following: unemployment duration, job search intensity, chance of quitting the job search, wages and hours immediately following re-employment, and the trajectories of wages and hours in the long-term following re-employment. Using the individual level panel data provided by the SIPP has the advantage of allowing us to follow individuals over long periods of time, enabling us to track trajectories for more than 2 years after a spell has ended and determine how minimum wages affect employment matching. In this way we can distinguish whether individuals are on average affected only in their job search, or if the effects of a minimum wage linger after re-employment.

The minimum wage for an individual spell is not necessarily constant over the spell duration, for this reason we distinguish two components of the minimum wage: its initial level and within spell variation (nominal). Separating these two features of the minimum wage policy allow us to better determine the effect of minimum wage levels on outcomes for the unemployed, as we isolate it from situations in which the minimum wage is raised. Furthermore, we can also determine how these increases of the minimum wage impact the unemployed, if at all.

We find overall that a higher initial minimum wage level has significant impact on search behavior only. On average, individuals who find themselves beginning their unemployment under a higher minimum wage level are more likely to quit their search and exhibit lower search intensity. However there is no significant impact on unemployment duration or starting wages and hours upon re-employment, we do not even find a significant impact on hours and wages trajectories over the next couple of years. We also find no evidence of skill based heterogeneity on these results.

Within spell variation of the minimum wage displays a stronger and more significant effect. We find that unemployed individuals who experience a minimum wage hike during their unemployment have on average 12% longer spells as well as lower search intensity and higher probability of quitting. In fact the effect on search behavior (quitting and intensity) is several times larger than it was for levels of the minimum wage. We also find that the effect on unemployment duration is heterogeneous by productivity, with the less productive workers being potentially the only significantly affected[1]. However, we find very little evidence that any of the impact of the minimum wage lingers on. Other than a potential reduction in working hours' trajectories, it seems that the policy has next to no effect on the wages or incomes of the unemployed (the effect on hours is suggestive but too weak).

As mentioned, surprisingly little attention has been afforded to unemployed workers in the minimum wage literature. Pedace and Rohn (2011) is the only other paper we are aware of that measures the impact of higher minimum wages on unemployed workers. Using a hazard model under four different distributions of the survival function and a sample of unemployed individuals from the Displaced Worker Survey, they study the impact of minimum wage levels on unemployment duration. They find an effect that varies along the lines of sex and educational attainment. Specifically, spells become longer for older women in low skill occupations and men that are high school dropouts; while more educated men reduced the length of their unemployment period. However, their analysis focuses exclusively on unemployment duration, and so they cannot conclusively say whether a higher minimum wage is beneficial or not for less educated women or more educated men, as longer spells are not necessarily a bad thing (they can be good if they lead to better wages for example). Additionally, our analysis reveals the existence of two different effects acting on the unemployed that are not distinguished by Pedace and Rohn. Summing these two effects makes the interpretation of their results more difficult and may contribute to the high heterogeneity they find.

Clemens and Wither (2019) is one of the most recent and relevant papers that evaluates the impact of the minimum wage using individual level panel data. Furthermore, the authors use the SIPP survey in their analysis too, albeit only the 2008 panel, and show that employment suffers when minimum wages are higher. However, they do not concentrate on the impact of

[1] In our analysis we proxy for low and high productivity workers by educational attainment, More details are given in the data and results sections.

higher minimum wages on the unemployed but rather on workers in general. Furthermore, they concentrate on a very particular moment in history in which two rare events met: the Great Recession and a 41% increase of the federal minimum wage. Our analysis seeks to provide policy relevant information in a much wider set of scenarios.

Previous research that has recognized the need for individual level panel data has typically reached similar conclusions to those of Clemens and Wither using other sources. Neumark and Wascher (1995) and Abowd et al. (1997) find dis-employment effects using the Current Population Survey (CPS), and Currie and Fallick (1993) find dis-employment effects as well using the National Longitudinal Survey of Youth (NLSY). However, Zavodny (2000) uses the CPS survey and finds no dis-employment effects for low productivity teenagers compared to other low productivity teens.

We improve on these analyses with the use of the SIPP survey. The SIPP data has two important advantages over the CPS data: it follows individuals for a longer term, four years compared to only one; and it provides continuous and high frequency data on them for the whole period. Compared to the NLSY its key advantage is that it not only follows young individuals but a representative sample of workers of all ages. Although teenagers have been a popular subject in this literature, this is a compromise that is no longer necessary with many current datasets.

An important body of research finds that minimum wages have a negative impact on employment (Powell, 2017; Meer and West, 2016; Thompson, 2009; Neumark et al., 2004; and see Neumark and Wascher, 2008 for a comprehensive review of previous work). However, another strand of this research questions the methods that lead to these results and find no dis-employment effects (Dube et al., 2010; Allegretto et al., 2011; Giuliano, 2013; Card and Krueger, 1993) or even positive effects (Card, 1992). We contribute to this debate using data on unemployed, who have not been studied so extensively even though they can provide some advantages. First, our results are robust to census divisions time fixed effects, unlike more aggregated data. Second, recent research has suggested that the effects of the minimum wage are felt on employment growth (Meer and West, 2016) and flows (Dube et al., 2016) rather than in levels. Under these circumstances, the unemployed are more likely to exhibit the effects sooner, if there are any. Furthermore, their outcomes should be less affected by sticky prices than the employed population (Barattieri et al., 2014).

Finally, the high frequency of the SIPP data allowed us to distinguish minimum wage levels from its hikes during a spell. We find that the effects from levels and changes are in fact different, even in the long term. Previous literature has hinted at this result before, with findings that increases in prices due to a change in the minimum wage occurs shortly after the hike (Aaronson, 2001; Aaronson et al., 2008; Basker and Khan, 2016), and that firm exit and entry is accelerated as well by minimum wage increases (Aaronson et al., 2018). Meanwhile, the critique by Sorkin (2015), that short-term and long-term effects are typically misinterpreted in the literature, is only valid for horizons significantly longer than ours (inflation has little effect in a window of three years).

The rest of the paper is organized as follows: in Section 2 we describe the data used; in Section 3 we summarize the econometric models used, as well as their advantages and limitations; in Section 4 we present our findings and our understanding of what they mean; in Section 5 we cover various robustness checks we implemented; finally, in Section 6 we conclude and summarize the most important findings and their implications for the minimum wage policy.

1.1 Data

We use four panels of the Survey of Income and Program Participation (SIPP) in our analysis, obtaining a sample of 79,082 spells longer than a week for people no older than 65, of which

40,851 are observed till their end[2]. The first three panels, for the years 2001, 2004, and 2008, follow participants for up to four years; and the 2014 panel for a maximum of only two years. In addition to this survey data we use data on state and federal minimum wages, and on monthly state unemployment rates.

Using the survey, we define a spell as a period of unemployment that begins when a person first declares to be looking for work and ends when said individual finds a job. Therefore, individuals who cease to be employed are not defined as unemployed unless they declare to be looking for work and the spell will only start once they admit to this. However, this also means that a spell does not end merely because someone declares not to be looking any longer, even if the search is never resumed within sample.

The monthly data on state minimum wages comes from Vaghul and Zipperer (2016). During the time span covered by the SIPP data just described the minimum wage takes values between \$5.15 and \$10.50 and was changed for 233 month-state pairs within sample. Of these, 89 changes come from increases of the federal wage floor in 2007, 2008, and 2009; and almost 90% occurred either in January (107) or July (96). The average nominal increase was \$0.54, but the largest changes surpassed \$1. Later in the paper we will exploit the way minimum wages are changed to design robustness checks to help dissipate fears of bias in our findings.

To study how minimum wage levels affect several outcomes of the unemployed, we use a difference-in-difference model. Our analysis considers three key characteristics of an unemployment spell: length, effort, and success. Additionally, four re-employment outcomes: starting wages, starting hours, and unconditional trajectories for wages and hours. Length of unemployment is measured in weeks simply as number of weeks that the unemployment spell lasted. To measure the effort put in job seeking we kept track of whether an individual

[2]Spells of length one week may lead to biases for some outcomes. Indeed, including them makes point estimates larger.

declared to be in the labor force each week and created an index for the percentage of time (weeks) spent looking. For success we created a binary indicator that takes the value one if the individual did not declare to be looking the last eight weeks of his spell (regardless of whether he found employment at the end), and zero otherwise. Re-employment outcome variables are used as they are found in the sample or in natural logs, with imputations using declared income when hourly wage is not available but income is.

We consider some important re-employment outcomes. Starting wage and hours after the spell is over, that is available for most uncensored spells in the sample, will allow us to discern whether the new wage floor is conducive to better or worse matches for the unemployed. Additionally, we keep track of these individuals' longer-term wages and hours. We record their wages and hours in intervals of four weeks after their spell is over, for more than two years (128 weeks). This allows to establish longer term effects of the policy, if there is any. Long-term outcomes make it possible to get a better estimate of the benefit or cost the policy has on the population it is imposed to. For example, a new minimum wage may mean a higher wage right after employment, but a worse wage trajectory, which we will be able to see.

However, it is also important to note that the sample shrinks very fast when we look at wages' and hours' trajectories, which has two important effects. The first, we may be concerned on the selection of those that disappear or remain in the sample. Most of them will be observations that end with the end of the survey itself, that only follows people for four years; we do not expect that these people should differ from other individuals in important ways. However, another group may be people whose spells were short (long), and so we get to observe their wages for longer (shorter) periods, if these people's wages differ importantly, then we may get biased estimates in our regressions. To test this, we did a correlation analysis between starting wage, after the unemployment period, and unemployment duration and found no correlation between these two outcomes, with or without controls. Nevertheless, we did find a correlation between unemployment duration and starting hours after the spell, but the coefficient is noticeably small in magnitude. An increase in 1% on unemployment duration produces an increase of only 0.05% in starting hours. An issue that remains is the loss of observations, and precision, over time. The analysis of future wages for example, starts with a little over 30,000 observations, but has lost about half by month seven (28 weeks), reaches 10,000 by month sixteen, and is below 2,500 for month thirty-two. This is particularly important for analyses of subpopulation, that can be small from the beginning, making for very imprecise trajectories in some cases. Although inference is still possible in most cases, we will not focus on any one estimate but instead only make general conclusions about the larger trends.

As a result of having weekly information on specific individuals, we can identify very closely how each spell responds to minimum wages. We even identify how unemployment spells respond to *changes* in the minimum wage that occur after the spell starts but before it ends. This gives us the opportunity to estimate not only how minimum wage levels impact the unemployed, but also how, or whether, its changes affect them. Furthermore, to establish the effect that minimum wage levels have on unemployment spells with weekly individual data we need to separate levels from changes. However, this gives us the opportunity to study another interesting aspect of the policy and its effects on the labor market.

Since we are considering several different outcomes, we sometimes use different samples when it seems appropriate. For the length of unemployment spells for example we can use a larger sample because duration analyses can deal with censored data, and so we do not need to ignore spells that are not observed till their end. Nevertheless, all analyses ignore workers that move to another state, of which our sample includes 4,879 spells (5.7% of the sample). Although most of them moved before or after their spell, not while unemployed, working hours and wage trajectories could be affected by including movers in the analysis, and to make results more comparable across outcomes we decided to ignore them completely.

Additionally, we ignore 6,381 spells that only last a week. Although including these in the analyses does not compromise our conclusions, our definitions do not deal well with spells of this length. For example, because we do not observe intra-week variation the way we define spells means that search intensity is always one for spells of length one. This may not be an important problem for levels of the minimum wage, but it is important for the effect of minimum wage changes, and both analyses should not be done separately. All of the above means that for the duration analysis we consider all spells that last at least two weeks of workers that remained in the same state throughout the survey. In this case we chose to treat spells longer than 52 weeks as censored spells, they enter the analysis but as if we did not observe them beyond 52 weeks. For behavioral outcomes (search intensity and search abandonment) we considered a similar group but ignored censored spells under 50 weeks[3], reducing the sample by an additional 31,659 spells. Finally, for re-employment outcomes we only consider uncensored spells not longer than a year (52 weeks) with information on next starting wage and hours, which leads to the additional loss of 5,100 spells, about 14%.

1.2 Model

We identify both the short-term effect of increasing the minimum wage on the currently unemployed as well as longer-term effects on their labor market outcomes. Short term effects are impacts of the policy change on spell length, search intensity, and abandoning the search; while the longer term outcomes that we analyze are the impacts on future wages and working hours. For unemployment duration analysis we use a hazard model that takes better advantage of the data than a linear model. For workers getting discouraged from their search (a binary outcome), we use a binary logit model. All other outcomes are studied using a linear model.

[3] Spells can be sensored at any time if it coincides with the end of the survey window.

1.2.1 Accelerated Failure Time Model

To determine the effect of treatment on spell length, we use an accelerated failure time model. We chose this model over a cox proportional hazard because the proportionality assumption, that requires the hazard ratio to be constant, is not met by the data. Instead we chose an accelerated failure time under a lognormal distribution because the proportionality tests show that a proportional hazard would underestimate the hazard at low durations and overestimate it for longer spells; a lognormal distribution should fit perfectly such data. With it we can establish how levels and increases of the minimum wage reduce or increase the probability that an unemployed worker at time t finds employment at time $t+1$. In our regressions we add several controls as well as the two variables of interest: initial real minimum wage, to capture the effect of the level, and actual real minimum wage in the cases where a change occurred during the spell, to capture the effect of changes. The hazard model takes the following form:

$$\ln(T_{iswt}) = \beta \times \ln(imw_{ist}) + \delta \times \Delta mw_{ist} + \gamma X_{iswt} + \lambda_s + \sigma_t + \epsilon_{iswt} \tag{1.1}$$

Where $log(T_{iswt})$ is the log of the failure time (spell length) for individual spell i in state s and week w of month t (identified uniquely for each year), imw_i is the initial value of the minimum wage at the start of the spell i, and mw_{st} is the actual value of the minimum wage. Δmw_{ist} is the difference of the natural log of the current minimum wage minus the initial minimum wage of spell i. This model includes state (λ_s) and month-year (σ_t) fixed effects, as well as individual covariates (X_{iswt}): age and its square, sex, race, number of children under 18 in the household, current unemployment rate (by month by state), educational attainment, and the reason they left their last job.

Traditionally for this literature we would have β be our parameter of interest, the impact of the minimum wage level on unemployment duration. However, with minimum wages changing during some unemployment spells but not others, the parameter β would be confounded by the effect of these increases, if they have any. Previous research (referred to in the introduction) hints that changes of the minimum wage may in fact have an important short term effect, which would mean that it is likely that this effect is not insignificant and that its sign and size matter. Therefore we separate initial level of the minimum wage from changes in it during a spell of unemployment, that in the regression accompanies δ. As a result, we get two parameters of interest, β and δ. The former informs us of the impact of a higher minimum wage on spell duration; and the latter about how these outcomes are impacted by being unemployed when the minimum wage is increased.

1.2.2 Linear Model

Most other outcomes we study can be evaluated using a standard linear model. We use this linear model for search intensity, log of starting wages, log of starting work hours, trajectories of wages (attributing a wage zero to those not working), and trajectories of working hours (attributing zero hours to those not working). However, in the case of whether a worker abandoned the search for employment (defined as not looking for employment the last eight weeks observed) we instead use a logistic model, given that the outcome is binary. Nevertheless, the covariates used are very similar to those used in the duration analysis, although in this case we cannot use real time values and instead choose to keep initial values of the covariates. The resulting equation is:

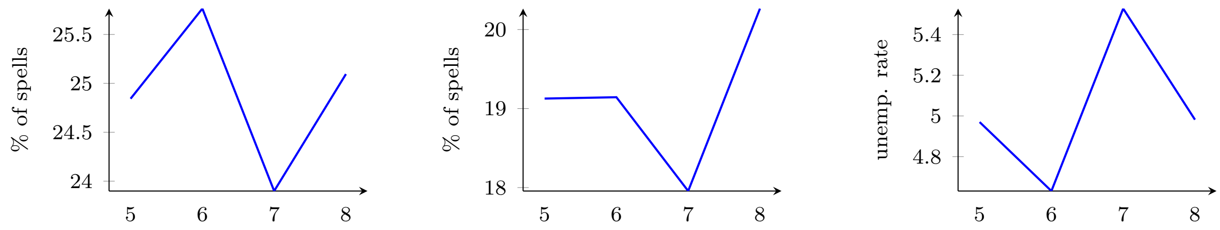
$$y_{ist} = \alpha + \beta \times \ln\left(imw_{ist}\right) + \delta \times \Delta mw_{st} + \gamma X_{ist} + \lambda_s + \sigma_t + \epsilon_{ist} \tag{1.2}$$

Where y_{ist} is the outcome of interest (search intensity, future wage, and future hours). There are some differences worth explaining between this model and the hazard model. All the covariates in this model correspond to their value when the spell started instead of their actual value at each point in time. Only one exception is made for the analysis of trajectories, in which we include both the initial and current unemployment rate by month by state, as a way to describe the individual (initial value) and the current macroeconomic conditions that are being endured (current value). On other respects the models are mirroring each other. For analyzing the outcome binary outcome "quit" we use a logit model with the same covariates.

1.2.3 Sample Balance - Minimum Wage levels

There are situations in which a difference-in-difference strategy may be unable to identify the proper effect of minimum wage levels. We are particularly concerned with the effect a higher minimum wage has on the composition of the unemployed. Research on minimum wages commonly finds some important effect that may lead to a compositional change of the unemployed under different minimum wage levels. Such a compositional change of the unemployed could compromise our results, as they might instead arise from comparing two different groups and not be informative of the effects of the policy. The data show that at higher minimum wage levels the labor force becomes more likely to have education beyond high school, and less likely to be in its teenager years. However, this can also be said of time, especially with the Great Recession in the middle of our sample. To limit the effect of time we graphed some of the most important characteristics of the unemployed only for panels 2001 and 2004, which do not include the federal minimum wage increases or the financial crisis, while still encompassing eight years. Figure 1.1 shows how education, preponderance of teenagers, and unemployment rate at the beginning of the spell, change with different minimum wages, pooled at their closest round number to create this graph. We observe no

Figure 1.1: Some Trends by Minimum Wage level - Panels 2001-4



(a) At most High School (b) Teenagers (c) Unemployment Rate

trend for any of these three characteristics.

To corroborate what is observed in these graphs we additionally regressed the minimum wage levels on these characteristics, adding time and state dummies. The results are reported in table 1.1, in which each column considers different samples that we will use for different outcomes. The lack of correlation between minimum wage levels and educational attainment, being a teenager, or the unemployment rate at the beginning of the spell is telling evidence that the unemployed should not be different at different levels of the minimum wage.

Even with no evidence that the sample's composition is changing due to different minimum wage levels, there can be differences between high and low minimum wage places that entail a problem for a difference-in-difference analysis. Allegretto et al. (2011) and Dube et al. (2010) make the argument that for any state a proper control state should come from within the same geographical division, because of important differences in labor market dynamics across U.S. census divisions. Consequently, we tested the robustness of our results to including division specific time trends. Another concern with this approach is that minimum wage changes may be correlated to certain local market conditions. In table 1.1 we show that there is no observable correlation between minimum wage levels and unemployment rate at the time spells start. We also address this issue in a robustness check that exploits federal minimum wage changes in 2007, 2008, and 2009, which are not driven by local conditions.

Table 1.1: Correlation Analysis for Minimum Wage levels

	Hazard (1)	Behavioral (2)	Wages/Hours (3)
Education $\leq$ High school	-0.00002	-0.0004	0.0001
	(0.0006)	(0.001)	(0.001)
Teenager	-0.0002	-0.001	0.0001
	(0.0006)	(0.0008)	(0.001)
Unemployment rate	-0.858	-0.783	-0.745
	(0.720)	(0.679)	(0.660)
Fired	0.002	0.002	0.001
	(0.001)	(0.002)	(0.002)
Quit	0.003**	0.004**	0.002
	(0.001)	(0.001)	(0.001)
N	80,586	45,754	33,195

significant at: *** 0.1% ** 1% * 5% + 10%

1.2.4 Sample Balance - Minimum Wage changes

For a subset of spells in our sample, a minimum wage change occurred while the worker was looking for employment. We observe 3,018 uncensored spells that experience this occurrence, representing 8.2% of uncensored spells (that are between two and fifty-two weeks, for non-movers).

We showed above that there is little evidence to suggest that compositional changes are occurring, validating the use of a difference-in-difference model. Further, for minimum wage changes we can argue to have a better comparison group, because in this case we also compare workers to others in the same state, that found work right before the policy change. The comparison group should then be even more similar for this analysis than for the previous one.

However, we might be concerned that changing minimum wage leads to other compositional changes in the sample of unemployed. We can deal with states having certain preference for the policy, or macroeconomic conditions being correlated to it, in the same way for changes in minimum wage as we do for levels, but other choices are important in this case that did not matter before. Minimum wage changes are typically known in advance and can be anticipated. We may have a problem if individuals or firms adjust in response to a minimum wage change that is about to happen. Table 1.2 compares four different groups we might be concerned about: those with spell length of one and those that moved to a different state, who we exclude from the analysis; those that do not experience a variation of the minimum wage while unemployed, and those that do. Important to highlight, we only see small differences in average wages, and almost no difference in unemployment rates when spells start. The latter observation being of interest because it contradicts the notion that minimum wage changes may be correlated to macroeconomic conditions, to which we alluded above.

Table 1.2: Summary Statistics

	Spell length=1		Movers		$\Delta mw = 0$		$\Delta mw \neq 0$	
	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.
Age	32.0	12.4	31.2	11.3	32.2	12.8	32.6	13.1
Women	0.439	0.496	0.508	0.500	0.460	0.498	0.466	0.499
0 to HS	0.197	0.398	0.150	0.358	0.229	0.420	0.237	0.425
Teenager	0.153	0.360	0.116	0.321	0.174	0.379	0.180	0.385
Race:								
White	0.778	-	0.751	-	0.734	-	0.703	-
Black	0.141	-	0.155	-	0.181	-	0.202	-
Asian	0.029	-	0.033	-	0.033	-	0.033	-
Other	0.053	-	0.061	-	0.052	-	0.062	-
Unemployment	0.062	0.021	0.062	0.022	0.065	0.022	0.065	0.021
Fired	0.032	0.177	0.030	0.169	0.033	0.178	0.034	0.181
Search Intensity	1	-	0.867	0.268	0.851	0.287	0.805	0.305
Spell length	1	-	16.1	19.1	15.0	11.5	25.1	13.3
Δmw	0.03	0.59	0.70	6.19	0	-	8.66	6.21
Observations	6,343		5,423		68,576		6,783	
Last wage	\$ 11.70	\$ 13.01	\$ 15.95	\$ 53.79	\$ 13.68	\$ 22.97	\$ 14.86	\$ 26.54
Observations	3,903		2,661		27,367		2,355	

Migration could be a response to changes in the wage floor. When minimum wage is increased in a state, a worker may choose to look for employment in another state as a response. In our sample of uncensored spells 4,471 unemployed workers moved to a different state, with

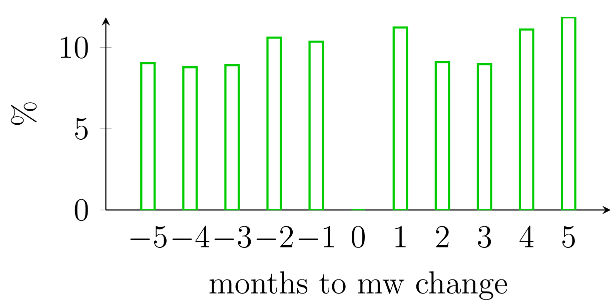
947 of them moving while unemployed. If the worker left because of the increased minimum wage, the group migrating may be a less (more) productive one than the average. If this is the case, ignoring them will likely produce shorter (longer) spell lengths as a response to the minimum wage hike. However, even though the bias could theoretically go in either direction, we would expect that, since we are talking of a higher minimum wage, the migrating worker *because* of this new wage floor is most likely the less productive one. Further, this worker would be looking specifically for lower minimum wages, in order to improve his chances of getting employed. To understand if this is the reason for these workers to be moving, we studied how the new minimum wage for these individuals related to the wage floor in their home state, and found no evidence supporting this theory. In 215 cases the receiving state had a lower minimum wage, and in 251 cases a higher minimum wage, leaving 481 migrations that resulted in no change in the minimum wage that the individual experienced. Overall, a very balanced situation with even a small bias toward movements to higher minimum wage states, which is not consistent with migration motivated by increases in the wage floor. Nevertheless, they are significantly more likely to experience a minimum wage change than the rest of the population, which led us to decide to discard movers from our analysis, even if the movement did not happen during the spell. Although they do not seem to be reacting to minimum wage changes, they many times create a change in minimum wage that is not comparable in nature to the change experienced by other workers.

Another important potential source of selection into or out of the sample comes from personal choices, both by firms and workers. To evaluate whether firms or workers are induced by minimum wage changes to modify their behavior in any way that may impact our sample we investigated spells that started or ended around a minimum wage change. Changes in minimum wages are known in advance by both parties, which makes it possible for either of them to anticipate the increase and modify their behavior in potentially important ways. If firms were to lay-off low productivity workers right before the change in policy for example, our sample of workers affected by the policy change would be unbalanced, and our results likely biased. Workers could make decisions that may unbalance our sample too. In their case, it may lead workers to increase their efforts in order to start a job before the minimum wage is increased, if they believe either that the market will be less dynamic after, or that they will benefit from the increase by finding employment before it happens.

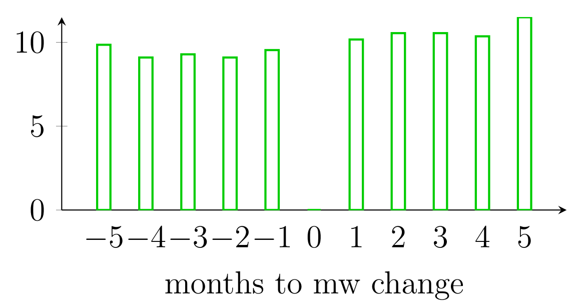
In order to test whether firms or workers were partaking in this type of behaviors we graphed spells starting and ending within five months of a minimum wage increase (figure 1.2). We only considered changes in January, or July for the years 2007, 2008, and 2009; but overall almost 90% of all changes in the wage floor. The figures show in green bars the proportion of spells started each month from five months before the wage floor increase to five months after.

Figure 1.2a show spells that start within the eleven months window. It seems most spells are starting after the policy is changed rather than before. We do not observe any sharp increase of spells in the months prior to the change either, which suggest to us that firms are not anticipating it by laying off workers. Next to it, figure 1.2b, similarly graphs spells that ended in the same time frame. In this case there are even less noticeable differences in the bars prior to the minimum wage being changed, that means no evidence that workers are increasing their employment rates in advance of the minimum wage change.

Figure 1.2: Spell started or ended at ±5 months of Minimum Wage changes



(a) Spell started



(b) Spell ended

1.3 Results

Our results deliver three main takeaways: (1) minimum wage changes have an effect of their own on the unemployed, (2) minimum wages have at best null effects for the unemployed on average, and (3) the negative effects we find are stronger for the least productive workers. We organized the discussion of our results in order to show these three takeaways clearly. We start by discussing overall effects and after that delve into heterogenous effects, concentrating on the populations that the policy seeks to protect.

1.3.1 Overall Effects

The most novel part of our analysis is the separation of minimum wage levels and changes, which means we allow two different effects on the unemployed. This was necessary in order to properly identify the effect of minimum wage levels technically, but it was also warranted by previous literature that has suggested that increarses of the minimum wage contribute an important shock to the market. Table 1.3 and figure 1.4 shows these effects for the unemployed, for all our outcomes.

A quick look at table 1.3 shows that in fact minimum wage levels and changes affect the unemployed differently in important ways. First, although the general direction of the effects is the same, the impact of a minimum wage increase on search outcomes is several times stronger than that of minimum wage levels. For unemployment duration, search intensity, and probability of discouragement we find coefficients for minimum wage changes that are several times larger than those of minimum wage levels (respectively eight, seven, and five times larger). Additionally, minimum wage increases make the unemployment spell longer, but higher levels of the minimum wage do not have a significant effect in durations. Second, we find suggestive evidence that minimum wage increases lead to a reduction in working hours over the following two years. However, we also find that the minimum wage has very little effect on re-employment outcomes, short-term and long-term, other than the above mentioned working hours. Nevertheless, a small negative impact is worse than it may at first seem, given that we would actually expect a positive impact from such a policy (i.e. the distance to the expected effect is larger than the distance to zero).

Table 1.3 displays coefficients for both of the effects we study, for the five short-term outcomes we consider. Column (1) shows that the minimum wage makes unemployment duration longer, but only when it varies during the unemployment period. The coefficient for Δmw implies that increasing the minimum wage by 10% would make employment the next week 12% less likely for an unemployed worker. Columns (2) and (3) are devoted to the search behavior outcomes, how intensely a worker looked and whether he got discouraged while looking. Both features of the minimum wage show a negative impact on these behaviors, although the effect is several times stronger for minimum wage changes 'pound for pound'. Finally, the last two columns (4) and (5) show how the minimum wage impacts short-term re-employment outcomes for these workers, finding that only working hours at next job are reduced by higher minimum wages and only about 2.5% for a minimum wage difference of 10%.

We think in particular the effect on unemployment duration is consistent with the notion that increases of the minimum wage act like a shock on the labor market, reducing demand for labor momentarily until the full effects of the new price are well understood and empirically observed. This could explain why unemployment spells are longer when minimum wages are increased during someone's unemployment period but are not so affected when the minimum wage is just higher. Figure 1.3 displays graphically what this 12% reduction in probability of finding employment translates to in terms of unemployment numbers. Assuming no more entrance or exit to unemployment other than employment (just to make the numbers as clear as possible), we see that after three months we can have more than 4% more unemployed workers just because we increased the minimum wage by $1. In July 2009 for example, when the federal minimum wage was raised seventy cents from $6.55 to $7.25 for two thirds of Americans and the unemployment rate was 9.4% this result would predict that just the minimum wage increase would lead to a unemployment rate 0.21 percentage points higher by the end of October (over 300 thousand more American workers unemployed).

We believe this effects originates on a demand shock, so we ran a few tests to see whether people that narrowly missed the minimum wage increases are similarly impacted. If the labor market is shocked by minimum wage increases as we presumed, we might expect that a worker that started looking shortly after a minimum wage change would endure a longer unemployment period as well. Although it is a flawed test, and only considers new unemployment spells, we find that there is a significant positive effect for workers that started looking within two months of the minimum wage increase. At the same time we do not find any significant effect for workers whose unemployment spells ended right before the minimum wage change[4].

However, although higher minimum wages do not seem to make unemployment spells longer they do affect unemployed workers' behaviors in the same general direction as minimum wage increases. We observe a negative relationship between minimum wages and both search intensity and probability of abandoning the search. An increase in minimum wage of 10% leads to a reduction in intensity of almost 10% from its baseline: a worker that spent three months unemployed will look one full week less on average if the minimum wage is increased by 10% during his spell. At the same time, a 10 point differential in minimum wages only lowers search intensity by less than 1.5%, equivalent to about a day less of looking for work

[4]In our testing we 'moved' the minimum wage change to assign it to workers in four different groups: spells that ended within a month before of the minimum wage change, and those that started within a month, two months, and three months after a change. We find only a significant effect for the group that lost their job within two months of the change, making us think the effect is short lived as well. However, we are a little surprised there is not a significant effect for workers that started looking within a month after the change, although this could have to do with sample size (30% smaller than the sample for the second month).

for a three month spell. We observe the same dynamic reflected on probability of abandoning the search for work, the effect is more than five times stronger for minimum wage changes. The same 10% increase or differential will lead to a likelihood of quitting the search larger by 2.8 percentage points in response to increases, and to 0.5 percentage points in response to a similar differential in minimum wages. This is a very large effect when we consider the sample baseline of 10% for probability of discouragement.

A small puzzle with these findings is to see behaviors 'worsen' at higher minimum wages, even though unemployment durations are unaffected. We do not find any significant loses related to higher minimum wages in the short or long term that could help explain these responses either. The better wages may be encouraging more workers to look, than on average look less intensely and are less attached to the labor force, but we do not find evidence that minimum wage levels have compositional effects on the pool of unemployed workers either. Still another possibility is that higher minimum wages make it easier for firms and people to find a desirable match, and allow for these new behaviors. If this is the case, unemployment durations may not be significantly affected just because the impact is too small for us to measure it successfully.

Finally, table 1.3 presents the results for short-term re-employment outcomes. We see that starting wage at re-employment (column 5) is not significantly affected, not even by higher minimum wages. However, starting hours at the new jobs are reduced by higher minimum wages even though not significantly affected by changes of the minimum wage. A 10% differential of the minimum wage level lead workers to a 2.5% reduction in working hours according to our findings. Overall, considering the low statistical confidence on the coefficient for starting hours for minimum wage levels, we see little evidence that the minimum wage has any significant impact on short-term re-employment outcomes.

We do not want to ignore the role of time in our analysis. Starting wages may not be affected but wage trajectory may be improved or injured at the same time. Figure 1.4

shows wages and hours monthly responses to minimum wages for almost three years after the unemployment spell is over. These graphs show that there is no improvement over time for wages or hours either. In terms of minimum wage levels, we see a movement to positive coefficient for both wages and hours after two years, but they remain mostly not distinguishable from zero statistically even with 90% confidence intervals. For minimum wage changes, we see absolutely no significant effect on wages after re-employment, but we do observe that during the following two years working hours seem to display a more persistent negative and statistically significant effect. Although the coefficients include zero on their 90 percent confidence interval more often than not, the effect for some groups of workers, more sensitive to the minimum wage, might reveal a more significant negative effect.

Summarizing, we find that the evidence is strongly against the notion that the average unemployed worker benefits at all from the minimum wage policy. However, we recognize that the minimum wage policy is not meant to impact the average worker, but rather a subset of workers that are recognized as in need of help vis-a-vis their potential/actual employers. We will now concentrate on those workers that are more likely targeted by the minimum wage policy to see how they are affected by the policy during their unemployment. Nevertheless, we do get some insights that should remain true even for subpopulations. Minimum wage affects the unemployed more through within spell variation than through its initial levels, and it does not show strong persistence over time.

1.3.2 Heterogeneity

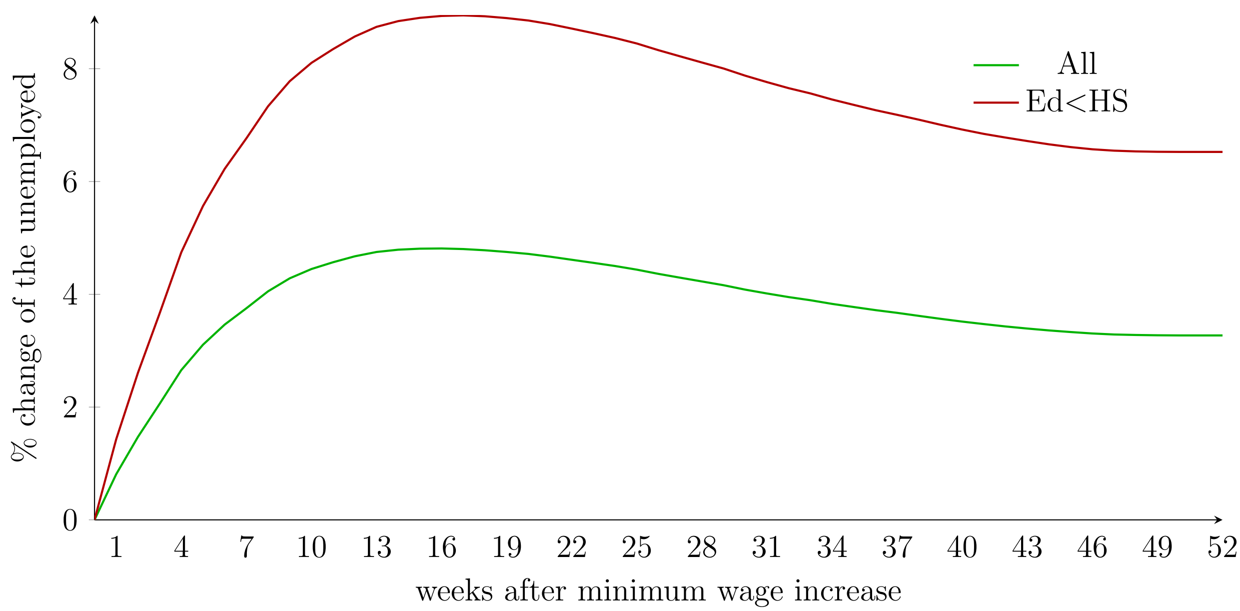
As mentioned, the minimum wage policy is not in place to help the average worker. It is only meant to improve the conditions of the worse-off workers, that may require some assistance to negotiate with their employers without being taken advantage of. To get a better look at the unemployed workers that should matter most for the minimum wage policy we considered

Table 1.3: Effect of Minimum Wage on short-term outcomes

	(1) Spell	(2) Intensity	(3) Quit	(4) Next Wage	(5) Next Hours
$\ln(imw)$	-0.154	-0.117***	1.998***	0.051	-0.247^{+}
	(0.129)	(0.031)	(0.469)	(0.056)	(0.124)
Δmw	1.191**	-0.765***	10.62***	-0.071	-0.414
	(0.368)	(0.122)	(1.131)	(0.088)	(0.267)
Intercept	2.963***	0.795***	-3.475***	1.006***	2.751***
	(0.224)	(0.050)	(0.922)	(0.125)	(0.219)
Observations	1,481,981	45,754	45,754	33,195	30,423

Significant at: *** 0.1% ** 1% * 5% $^{+}$ 10%. For spells we use an accelerated failure time model, for quitting we use a logit model, and for all other outcomes we use a linear model. All models control for state and time (month-year) fixed effects, education level, sex, race, age, age^2, unemployment rate, and number of children in the household. We exclude spells greater than 52 weeks, equal to one week, and spells associated with individuals who move. Standard errors are clustered at the state level

Figure 1.3: Increase in the Unemployed after a Minimum Wage increase of $1



several demographic categories. We broke out the sample by education: at most high school compared to more educated unemployed workers; by race: white, black, asian, and others; by age: teenagers compared to adult workers; and by information on their last wages by interacting the effects with the workers' logged last wage in constant dollars.

Although historical information on wages should be the ideal way to distinguish low productivity workers, there are issues with how this information is reported. Particularly important is the fact that over half the sample does not include information on past wages, but also important is the fear that the information may be given less accurately than other characteristics that are less likely to be plagued by error from the interviewee or the interviewer. Nevertheless, we used past wages distributions to decide which demographic characteristics separate workers more closely by their productivity. As expected, the group with the lowest average hourly wage is teenagers. Furthermore, their distribution is rather compact (50% earn between \$7.8 and \$10.4 in our sample) and the distance with adult workers is very significant (adults earn on average \$17.9 hourly, double the teens' rate). However, teenagers are problematic for more important issues in our view. They are much more likely to abandon the labor force in the next few years to continue their education than older workers, and are not the target population for the minimum wage policy.

We test racial disparity by using the four categories in the variable 'ERACE' (white only, black only, asian only, residual), but past wage distribution for these groups are significantly closer. Sample averages for each racial group range from \$14.1 to \$18.6, less than a third higher. The difference in average wages between women and men are similarly small, with an average wage of \$14.8 for the former and \$18.5 for the latter group. However, wage differential is significantly larger for unemployed workers by their educational attainment. Those that did not finish high school average \$11.9 while workers that advanced further average \$18.0, almost as large a difference as that between teenagers and adults. The dispersion for the less educated workers is also relatively small, 75% earn an hourly wage under \$12.8. For this reason we will concentrate on discussing in this section the results by educational attainment in which the low productivity workers are represented by respondants that did not finish high school.

Table 1.5 shows the result of interacting the minimum wage with educational attainment for our five short-term outcomes. We find that heterogeneity is important for three of these outcomes: unemployment duration, search intensity, and working hours at re-employment. According to this table unemployment duration is only affected for the less educated workers, and the probability of finding employment is reduced by 21% for these workers, almost double the finding presented previously. This is a critical consideration to evaluate the policy's impacts as according to the Bureau of Labor Statistics the unemployment rate for high school dropouts in july 2009, when the federal minimum wage was increased by $0.70 and the overall rate was 9.4%, was 15.3% (+63%). Our graph 1.3 depicts in the red line the effect for these workers, an increase of $1 on the minimum wage would lead to an increase in unemployed workers of 8.7% three months later. Therefore, our estimates suggest by October 2009 the unemployment rate would be 0.38 percentage points higher for high school dropouts as a result of the minimum wage increase in July (assuming they do not leave the labor force).

Columns (2) and (3) of table 1.5 reveal a rather surprising result for minimum wage increases when exploring whether heterogeneity by education is important for search behavior. The first rows show response to initial levels of the minimum wage and are not all that surprising. We see that search intensity goes down solely for workers that finished high school, even though neither group suffers from longer unemployment spells. At the same time, working hours at re-employment are also reduced only for the more educated workers, which could be the reason for their small drop in search intensity. Our analysis of longer term re-employment outcomes in figure 1.6 reveals no further heterogeneity or impact of the initial level of the minimum wage either. However, the results are rather surprising for minimum wage changes, the overall importance of the observed heterogeneity seems small, in particularly when considering the effect on unemployment duration for each group. These results for minimum wage increases raise the question of how can the two groups behave so similarly when one of them is suffering a lengthening of their unemployment spells and the other one is not. At least partially, the explanation may be found on our re-employment outcomes (columns 4 and 5). We see that working hours at re-employment are reduced significantly (both statistically and in terms of sheer size) only for the group with more education. This could make this group less driven to find employment, as they expect a lower revenue of doing so, at the same time high school dropouts could reduce their search efforts as a response of longer spells, even though their hours remain constant. Furthermore, figure 1.8 shows some evidence that the working hours response remains heteregeneous for a couple years against workers with high school diploma, which is also consistent with them showing negative behavioral responses.

Table 1.5 shows that there is some important heterogeneity on the effects ot the policy on unemployed workers. In particular, it shows that higher initial levels of the minimum wage have almost no effect on the less productive unemployed workers, as proxied by their educational attainment. This analysis suggest that minimum wage levels only play a small role in the unemployed workers' search behavior, for those with lower educational attainment. However, it also shows that the impact of minimum wage increases can be substantial. Unemployment rate can be increased significantly through these changes for the less productive workers. Furthermore, the impact on search behaviors is large for all workers and should be an important concern, as many seem to be pushed out of the labor force as a result of these increases.

Another interesting set of results is that by sex. Table 1.4 shows heterogeneous results for men and women. Initial levels of the minimum wage reduce unemployment duration for women, increase search intensity, increase probability of discouragement for women significantly less than for men, and reduces their wage at re-employment with respect to men (although not significantly from zero). With exception of the last effect on re-employment wages, higher minimum wages seem to benefit women relative to men. However, women fare much worse if their unemployment coincides with a minimum wage increase. Then their unemployment will not last longer than that of men (although it will be longer than without the minimum wage increase), but they will be significantly more likely to search less intensely, to abandon the search, and lose wages at re-employment (with respect to both men and zero). These effects are also orders of magnitude larger than those associated to initial levels of the minimum wage, the negative effect on wages for example is almost six times larger.

These results by sex are heterogeneous but not in the way results by educational attainment were heterogeneous, which is why they are interesting. Educational attainment by sex shows that in fact women are more likely to have advanced degrees than men, even though their wages are on average lower ($14.82 to $18.54). One potential explanation for these differences between men and women is labor force participation and attachment. The results in table 1.4 show that men are more likely to abandon their search in response to higher minimum wages, maybe this relatively lower attachment to the labor force allowed men to maintain their wages even under different demand conditions. Less competition by men might even help explain why unemployment durations are reduced for women. However, this explanation does not fit well with the findings for minimum wage increases. If the minimum wage is increased while unemployed, female workers become more likely to search less intensely and more likely to abandon the job search than men, not less. We have to keep in mind that now the whole market was shocked, labor demand will momentarily loose dynamism, and the worst conditions may explain why women are more affected by the increase of the minimum wage.

However, the heterogeneity by sex seems to be a short-term phenomenon. We do not see important differences in trajectories for either effect, wages and hours after re-employment

Table 1.4: Effect of Minimum Wage on various outcomes (Women)

	(1) Spell	(2) Intensity	(3) Quit	(4) Next Wage	(5) Next Hours
$\ln(imw)$	-0.0243	-0.144***	2.334***	0.0697	-0.188
	(0.135)	(0.033)	(0.477)	(0.054)	(0.135)
$\ln(imw) \times Women$	-0.274**	0.0562***	-0.628**	-0.0909**	0.124
	(0.105)	(0.012)	(0.222)	(0.029)	(0.072)
Δmw	1.262*	-0.656***	9,528***	0.170	-0.377
	(0.524)	(0.090)	(1.048)	(0.130)	(0.269)
$\Delta mw \times Women$	-0.158	-0.240*	2.315*	-0.537**	-0.0858
	(0.608)	(0.114)	(1.169)	(0.197)	(0.358)
Intercept	2.726***	0.842***	-4.038***	1.082***	2.643***
	(0.257)	(0.052)	(0.930)	(0.127)	(0.220)
Observations	1,481,981	45,754	45,754	33,195	30,423

Significant at: *** 0.1% ** 1% * 5% $^{+}$ 10%. For spells we use an accelerated failure time model, for quitting we use a logit model, and for all other outcomes we use a linear model. All models control for state and time (month-year) fixed effects, education level, sex, race, age, age^2, unemployment rate, and number of children in the household. We exclude spells greater than 52 weeks, equal to one week, and spells associated with individuals who move. Standard errors are clustered at the state level

seem to behave similarly and look generally unaffected by the minimum wage during the unemployment spell.

1.4 Robustness Checks

A very important discussion within the minimum wage literature refers to methods. We use a nation-wide approach that has been questioned by many important researchers as unable to account for local heterogeneity. Allegretto et al. (2011) show that disemployment findings were not robust to adding census division time fixed effects to the model most commonly used, and that we use in our analysis. Table 1.6 shows the effect of adding these fixed effects to each outcome's overall effect (column 1 of each table). We see that significant results remain significant and the same sign, and even some non-significant results gain significance for changes of minimum wage within the spell. Other responses to higher minimum wage

Table 1.5: Effect of Minimum Wage on various outcomes (Education)

	(1) Spell	(2) Intensity	(3) Quit	(4) Next Wage	(5) Next Hours
$\ln(imw)$	-0.120	-0.116***	1.850***	0.0519	-0.292*
	(0.140)	(0.031)	(0.466)	(0.058)	(0.126)
$\ln(imw)\times < HS$	-0.145	0.0399***	0.167	-0.0143	0.236**
	(0.110)	(0.019)	(0.269)	(0.053)	(0.077)
Δmw	0.738	-0.699***	10.61***	-0.0781	-0.613*
	(0.408)	(0.112)	(1.093)	(0.103)	(0.239)
$\Delta mw\times < HS$	2.104**	-0.134*	-1.282	-0.00865	1.070*
	(0.812)	(0.156)	(0.909)	(0.216)	(0.457)
Intercept	3.158***	0.645***	-2.435**	1.054***	2.848***
	(0.261)	(0.054)	(0.926)	(0.125)	(0.212)
Observations	1,481,981	45,754	45,754	33,195	30,423

Significant at: *** 0.1% ** 1% * 5% $^{+}$ 10%. For spells we use an accelerated failure time model, for quitting we use a logit model, and for all other outcomes we use a linear model. All models control for state and time (month-year) fixed effects, education level, sex, race, age, age^2, unemployment rate, and number of children in the household. We exclude spells greater than 52 weeks, equal to one week, and spells associated with individuals who move. Standard errors are clustered at the state level

levels more than double with respect to the original estimates and become significantly more imprecise, but remain statistically different from zero and keep their original sign. Meanwhile, the impact of these fixed effects on the responses to changes of the minimum wage while unemployed become stronger, not weaker.

We are also worried about correlation between local labor markets and minimum wage levels or changes, as well as the effect of differences in preferences for the policy that may lead to selection issues. Table 1.7 evaluate a potentially different effect from local state increases and general federal changes of the minimum wage. This table shows that federal minimum wage changes do not have a significantly different effect from other minimum wage changes, providing no evidence that this potential issue is causing any bias in our analysis. Therefore, local market conditions are likely not an important concern when interpreting our findings.

New minimum wage legislation typically comes into effect in January or July for our time period (almost 90% of all changes). In table 1.8 we explore the possibility that the effect may

Table 1.6: Regressions with Census Division time fixed effects

	(1) Spell	(2) Intensity	(3) Quit	(4) Next Wage	(5) Next Hours
real imw	0.280***				
	(0.072)				
ln(real imw)		0.442***	-11.59***	0.020	-0.098
		(0.097)	(2.927)	(0.075)	(0.119)
real mw	0.163***				
	(0.006)				
Δmw		-0.006***	0.085***	-0.002*	-0.005*
		(0.001)	(0.014)	(0.001)	(0.002)
Observations	74,611	47,561		35,461	32,289

Significant at: *** 0.1% ** 1% * 5% + 10%. For spells we use an accelerated failure time model, for quitting we use a logit model, and for all other outcomes we use a linear model. All models control for state and time (month-year) fixed effects, education level, sex, race, age, age^2, unemployment rate, and number of children in the household. We exclude spells greater than 52 weeks, equal to one week, and spells associated with individuals who move. Standard errors are clustered at the state level

Table 1.7: Regressions with Federal minimum wage change interaction

	(1) Spell	(2) Intensity	(3) Quit	(4) Next Wage	(5) Next Hours
Δmw	0.156***				
	(0.010)				
×Federal	-0.212				
	(0.139)				
Δmw		-0.007***	0.080***	-0.001	-0.004
		(0.001)	(0.010)	(0.001)	(0.003)
×Federal		0.003	-0.005	0.001	-0.001
		(0.002)	(0.026)	(0.004)	(0.006)
Observations	74,611	47,561		35,461	32,289

Significant at: *** 0.1% ** 1% * 5% + 10%. For spells we use an accelerated failure time model, for quitting we use a logit model, and for all other outcomes we use a linear model. All models control for state and time (month-year) fixed effects, education level, sex, race, age, age^2, unemployment rate, and number of children in the household. We exclude spells greater than 52 weeks, equal to one week, and spells associated with individuals who move. Standard errors are clustered at the state level

Table 1.8: Regressions for Spells started only in April (or October for 2007-2009)

	(1) Spell	(2) Intensity	(3) Quit	(4) Next Wage	(5) Next Hours
real mw	0.100***				
	(0.010)				
Δmw		-0.006***	0.062***	-0.003	0.007
		(0.001)	(0.018)	(0.006)	(0.005)
Observations	6,911	3,714		2,726	2,465

significant at: *** 0.1% ** 1% * 5% + 10%

be very different for spells that are started in months that would make them very unlikely to be affected by minimum wage variation while unemployed. This table only considers spells started in April for all years but 2007, 2008, and 2009, for which we only keep spells started in October, three months after the federal minimum wage change occurred. We can see that the effect of variation of the minimum while unemployed is unaffected by this test for all outcomes.

Finally, we addressed the possibility that our arbitrary choice of eight weeks not looking at the end of the observed spell may be driving the results on unemployed workers getting discouraged in their job search. Table 1.9 shows the results of choosing ten weeks instead of eight for initial minimum wage (minimum wage level), and table 1.10 shows the effect on the effect of minimum wage variation within the spell. We see in both cases only small changes in the coefficients, that however remain in other ways the same: same sign and overall meaning. We conclude from this that our choice of eight weeks is not indispensable to the finding and can be modified without need to modify our conclusions

We believe these robustness checks confirm that the effects we attribute to initial minimum wage and its within spell variation are causal. We find no evidence to suggest that these effects could be an artifact of the way we analyze the data, instead they seem to confirm that our findings are strong to alternative specifications.

Table 1.9: Regressions for Quitting Search defined around 10 weeks - Effects on Levels

	(1)	(2)	(3)	(4)	(5)
ln real imw	-5.226***	-5.124***	-4.692***	-5.323***	-5.243***
	(0.778)	(0.753)	(0.775)	(0.781)	(0.766)
ln(real imw)×:					
Education<High school		-0.174			
		(0.439)			
Women			-1.001**		
			(0.355)		
Teenager				0.322	
				(0.507)	
Black					0.145
					(0.510)
Asian					-1.416
					(1.026)
Other race					0.449
					(0.784)
Observations			47,561		

Table 1.10: Regressions for Quitting Search defined around 10 weeks - Effects on Changes

	(1)	(2)	(3)	(4)	(5)
Δmw	0.078***	0.078***	0.071***	0.075***	0.083***
	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)
$\Delta mw\times$:					
Education<High school		-0.009			
		(0.009)			
Women			0.015		
			(0.012)		
Teenager				0.013	
				(0.010)	
Black					-0.016
					(0.013)
Asian					-0.011
					(0.020)
Other race					-0.026
					(0.026)
Observations			47,561		

1.5 Conclusions

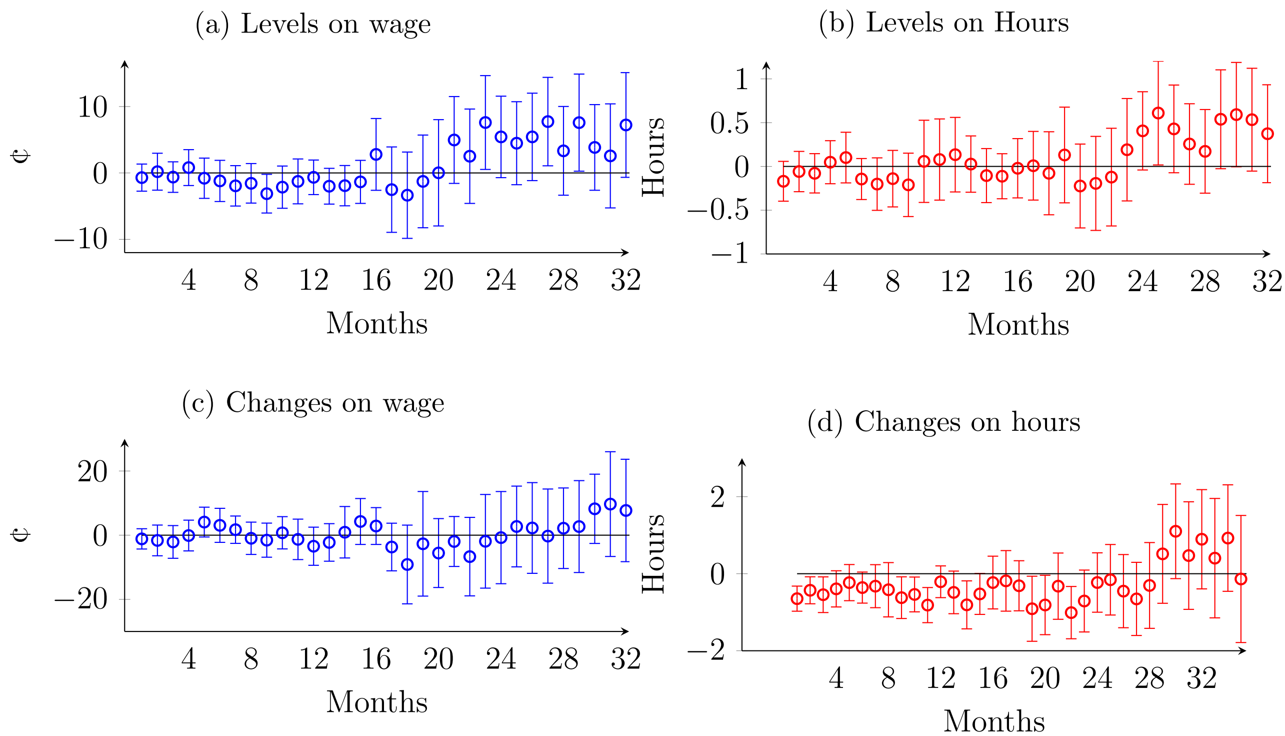
We examine how the minimum wage impacts the unemployed, both through its levels and its changes. We believe that the "levels" effect can be interpreted as the longer term impact of the minimum wage while the "changes" effect can be interpreted as the short term response to a shock to the market. We consider both how the policy impacts workers' unemployment outcomes and their longer-term re-employment outcomes. We chose to study the unemployed as we might expect them to be more vulnerable and exposed to the negative impacts of a policy implementation. In addition they are less bound by sticky wages. Thus if there are negative impacts of the minimum wage it is likely to be more pronounced and also more immediately observable for unemployed individuals. This allows us to more easily determine if negative effects exist, and also to investigate how a particular vulnerable group may be impacted.

We find that higher minimum wage levels are not creating gains for the unemployed workers in terms of any of the variables considered (spell duration, search intensity, quitting, wage and hours trajectories). We observe that on average individuals under a higher minimum wage level are searching less and quitting the search more, although the driving force behind this change in search behavior is unclear. We also find that with respect to levels there is much less heterogeneity between groups, with lower educated individuals being similarly affected as those with more education.

Interestingly, we find much stronger impacts of minimum wage hikes on those who are unemployed in the vicinity of a hike. Those unemployed during a change experience signficiantly longer spells, search less intensely and quit the search more. There is also very little evidence that it impacts the trajectory of wages and hours after re-employment. This provides some evidence that the negative impacts of the minimum wage are concentrated in the proximity of a change and do not resonate over a long time horizon. Previous research has made the point that all minimum wage effects are in fact short term effects because of inflation (Sorkin, 2015). Our findings suggest there is another reason for the short-term nature of minimum wage effects.

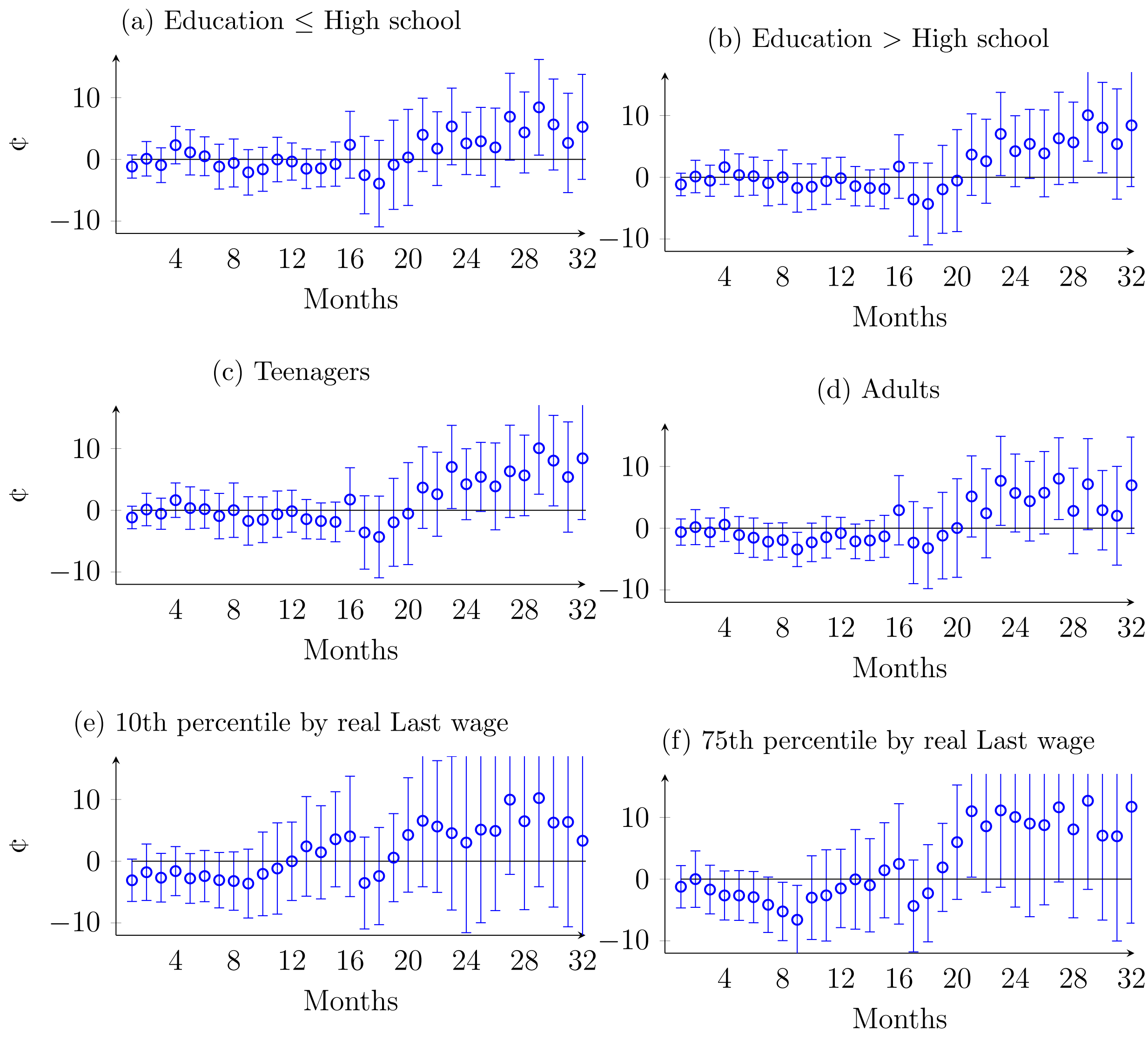
Overall, our results are critical of the minimum wage policy, giving minimal evidence of a positive impact to its target population, and strong evidence of negative effects on certain outomes (although mainly with regard to changes), in particular for less productive groups. However, our analysis focuses on a small section of the whole labor market, and an evaluation of the policy needs to consider the whole body of research devoted to it, in which we observe a variety of findings. A natural extension of this paper is to perform a similar analysis for those who are employed, and divide the effect as we have done to see how they are effected by the minimum wage. These results will likely not settle the debate on the minimum wage, but should raise some concerns about the policy. They also raise a concern with how the minimum wage is being increased today, with small changes every year. Should we care more about the impact of changing the minimum wage or about the impact of its level? Perhaps we are willing to accept short term losses in exchange for longer term benefit. Future research can take advantage of the increasing number of states indexing their minimum wages to inflation to provide a more definite answer to this question.

Figure 1.4: Overall effects on Wage and Hours trajectories



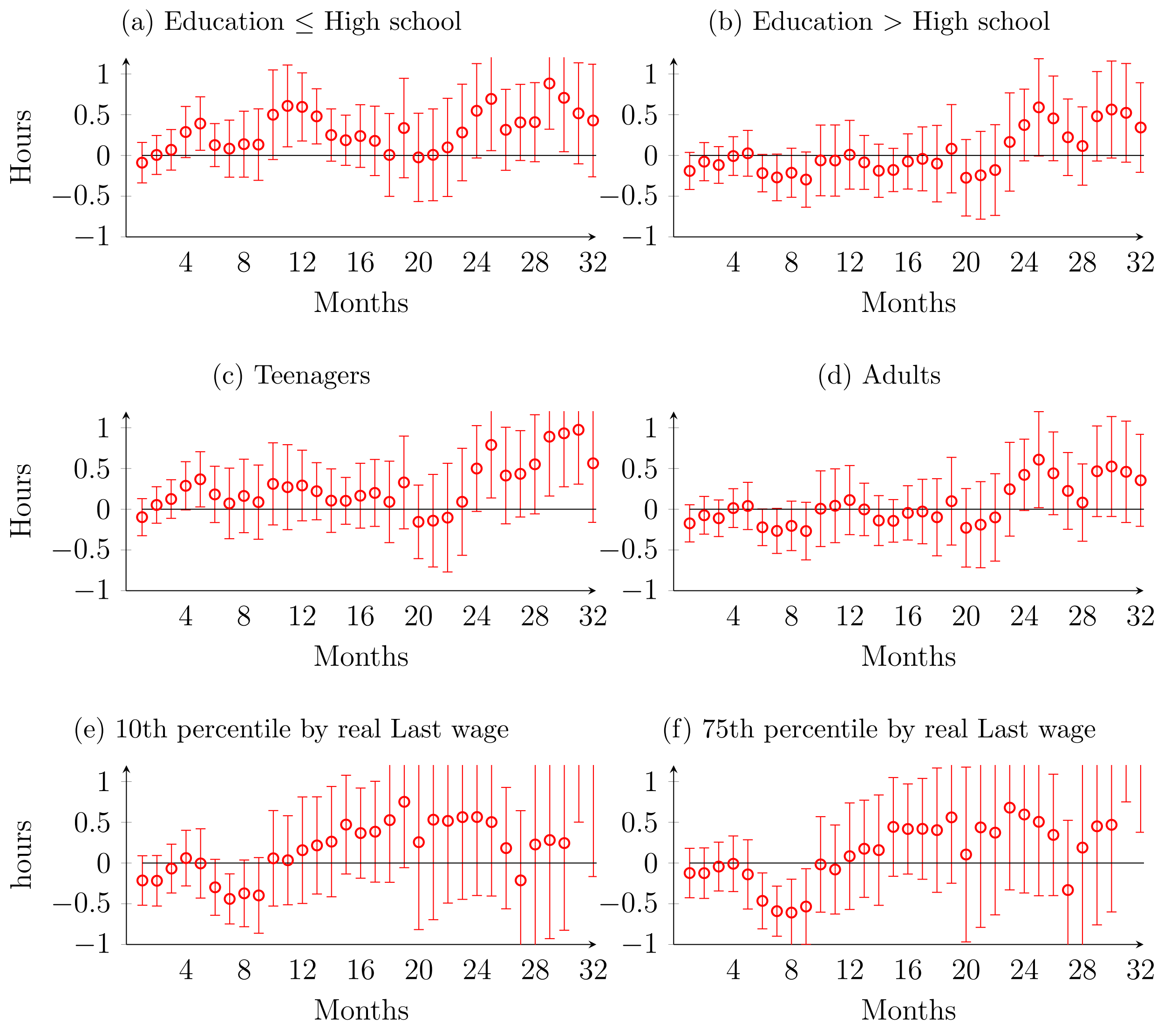
All graphs include 90% confidence intervals around the estimated coefficients.

Figure 1.5: Effect of a 1% Higher minimum wage on Wage's trajectories



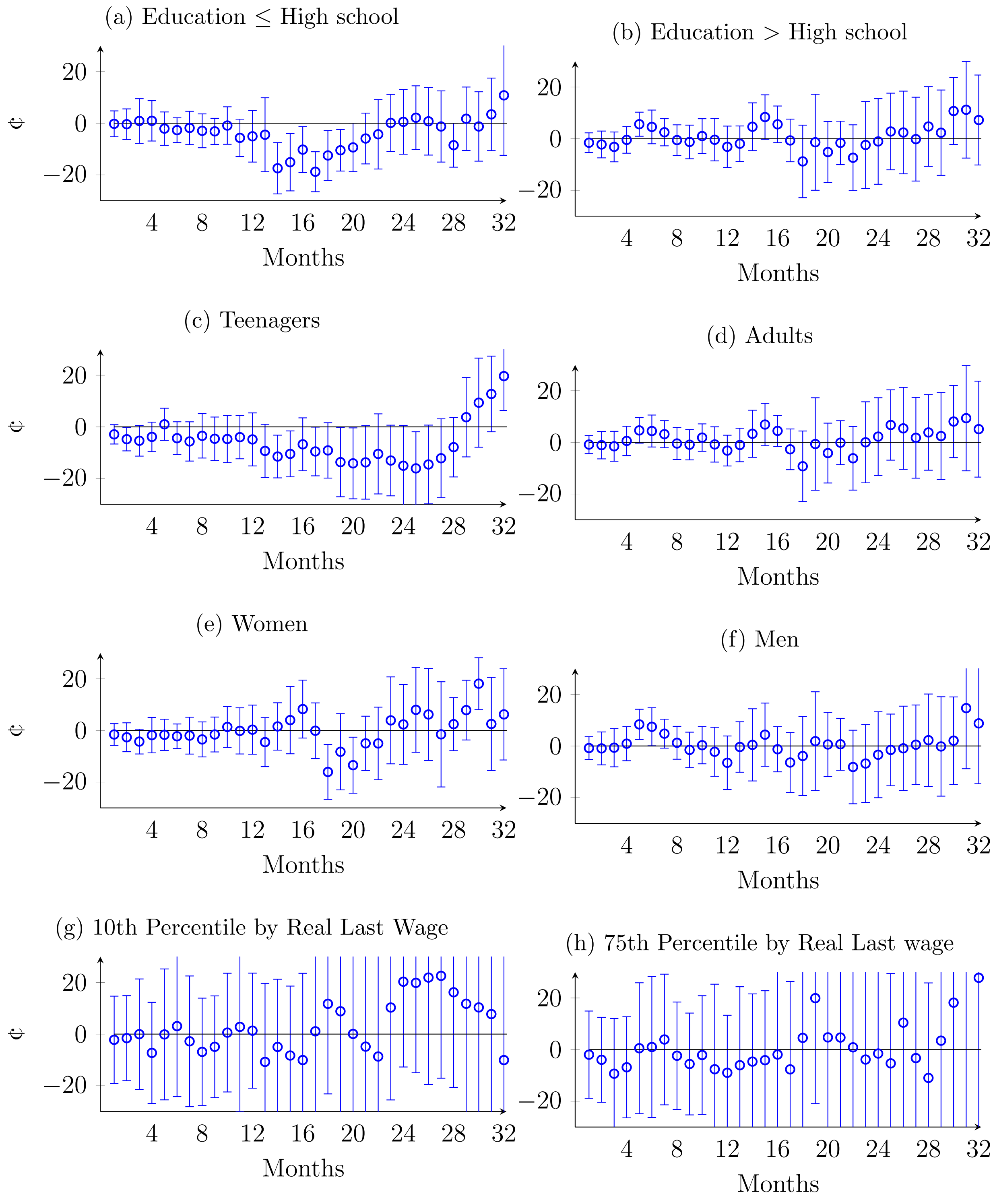
All graphs include 90% confidence intervals around the estimated coefficients.

Figure 1.6: Effect of a 1% Higher minimum wage on unconditional working Hours' trajectories



All graphs include 90% confidence intervals around the estimated coefficients.

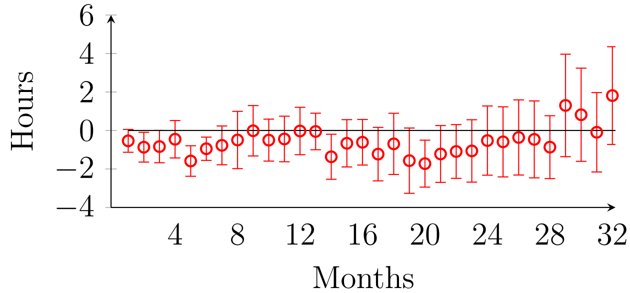
Figure 1.7: Effect of an Increase of the Minimum Wage on Wage trajectories



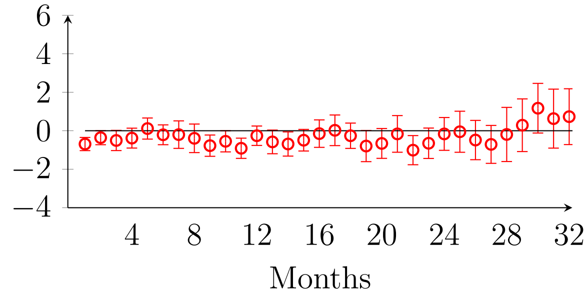
All graphs include 90% confidence intervals around the estimated coefficients.

Figure 1.8: Effect of an Increase of the Minimum Wage on Unconditional Working Hours' Trajectories

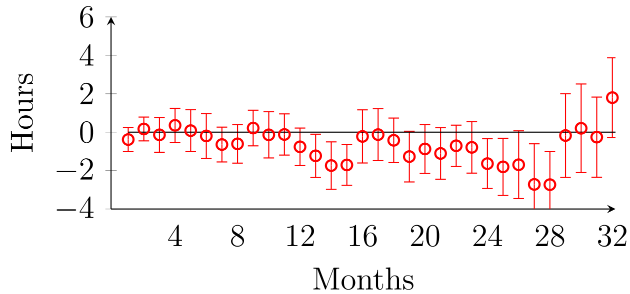
(a) Education ≤ High school



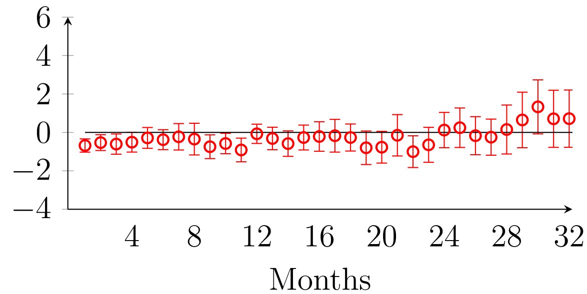
(b) Education > High school



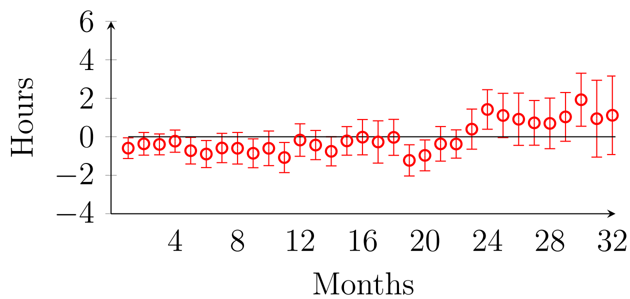
(c) Teenagers



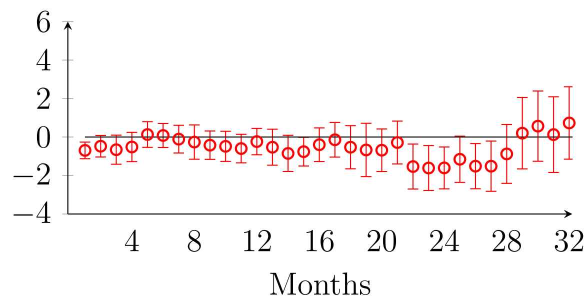
(d) Adults



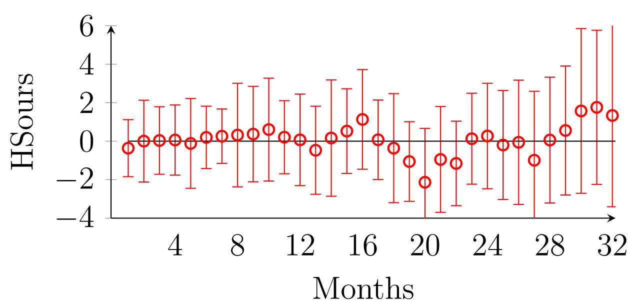
(e) Women



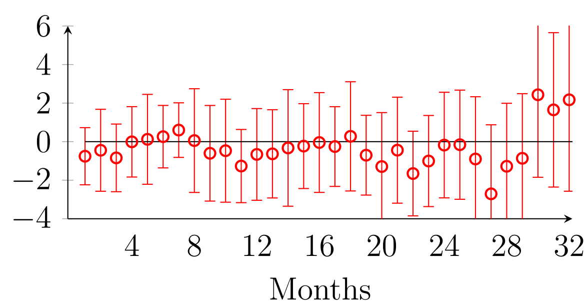
(f) Men



(g) 10th Percentile by Real Last Wage



(h) 75th Percentile by Real Last Wage



All graphs include 90% confidence intervals around the estimated coefficients.

Chapter 2

Mothers' labor supply and conditional cash transfers: Evidence from Chile

This paper represents the first effort to measure labor supply distortions created by Conditional Cash Transfer programs (CCT) for relatively schooled beneficiaries. CCTs became very popular at the end of the nineties. Starting with three programs in 1997, a report by the World Bank estimates that sixty countries had active CCT programs by 2018, with a combined one hundred and five million beneficiaries. The same report points to one likely reason for their popularity: CCTs seem to outperform all other types of social safety instruments in targeting the poorest households (The World Bank, 2018). Their simplicity also contributes to their popularity; even low income countries with limited state capacity can use them successfully.

In this paper, I focus on measuring the labor supply distortions created by a Chilean CCT program using a difference-in-difference strategy. The *Subsidio Unico Familiar* (SUF, 'unique family subsidy') was significantly expanded between 2007 and 2010, during which time it doubled its reach from one to two million people; at the same time, its monetary value increased by more than 50%. The SUF provides a cash transfer to families conditional on their children attending school and regular medical checkups. However, a potential unintended effect of these transfers could be to distort recipients' labor supply choices.

My analysis shows that certain groups of mothers do modify their labor supply as a response to the transfer by the Chilean SUF. Specifically, I find a positive labor force participation distortion for mothers in the age range 25 to 50 that are part of a couple, who now work more. On the other hand, young mothers in the age range 18 to 24 reduce their labor supply in response to the program. Finally, I observe a reduction in working hours for those that work, concentrated on younger and on less educated mothers.

The primary contribution of this paper is the analysis of a mature CCT program. At the time of its expansion, the SUF was 26 years old and the mothers benefiting from it had on average over eight years of schooling. In contrast, evaluations using randomized control trials are done at the birth of the CCT program (within the first five years), when benefited mothers have on average very little schooling. Because these are rapidly changing countries, we would be wrong to think that a relatively rich country today, such as Chile, would have nothing of value to offer on this matter. The fact is that Chile's per capita GDP by PPP in 2007 ($18,373) implies it is a better comparison for Mexico in 2017 ($17,331) than even Mexico itself in 2000 ($15,683). This is true not only for Mexico, but for many other countries that use CCT programs (Federal Reserve Bank of St. Louis, 2019).

If more education has an impact on labor supply choices and/or child investments, then we could reasonably expect labor distortions due to CCTs to differ at different levels of parental education. Indeed, there is evidence from prior research that suggest that more educated mothers will have fewer children and work more. Furthermore, this evidence also suggests that more educated mothers will invest more in health and education for their children. Over time education could have a critical impact on who receives the CCT, as well as on the effectiveness of the CCT to increase parental investment in children. These changes have potentially important implications for labor supply choices as well, and warrant research to measure the program's impact under these new conditions.

Measuring the labor supply distortions created by CCT programs is important because neither their sign nor size is easily deduced. Alzúa et al. (2013) suggest three main sources of distortions that could push mothers to change their labor supply choices, other than general equilibrium effects[1]. One potential source of distortion would have a negative sign, the pure income effect from non-earned income. It will theoretically lead to lower labor supply and should depend on the size of the transfer relative to the beneficiary's income. The other two distortionary forces would lead to an increase of mothers' labor supply. First, mothers may be able to work more now that they do not have to take care of their children, who are now at school. Second, if child labor is an important income source for the family, the condition imposed by the program will mean a decrease in the family's total income, forcing the mother to increase her labor supply to offset this effect.

Critically, as educational levels rise, these two last effects would become less important, while the first force needs not to be affected. Children are going to school now because their more educated parents want it, and so less time is freed from childcare and less income is lost from child labor. Therefore, labor supply distortions become more and more negative as the population becomes more educated, and estimates of labor supply distortions from less educated populations are not informative for more educated beneficiaries.

The rest of the paper is organized as follows. Section II discusses the current state of knowledge on the effects CCTs have on labor supply. Section III is devoted to explaining thoroughly the SUF subsidy and its recent evolution, including the event that will provide me with an identification strategy. Section IV describes the data used, Section V defines treatment and control and Section VI presents the difference-in-difference model. Section

[1] General equilibrium effects have a theoretical sign and magnitude that is ambiguous. However, given that the transfer under Chile's CCT is small relative to income, general equilibrium effects are likely to be of second-order importance in this case.

VII provides a discussion of the results obtained for several labor supply outcomes. Finally, Section VIII concludes and suggests ways forward.

2.1 Background

2.1.1 Literature Review: CCTs and Labor Supply distortions

Banerjee et al. (2017) review seven CCT evaluations using randomized controlled trials for four developing countries. They find no evidence that any of these programs affected labor supply negatively. However, this does not mean that the point estimates these paper find are ambiguous in sign or irrelevant in size. In fact, one of this studies finds effects on employment for Honduras, Nicaragua, and Mexico that, even though not statistically significant, respectively represent reductions from the baselines of 5.2%, 11.3%, and 5.1%[2] (Alzúa et al., 2013). Similarly, Skoufias and Di Maro (2008) find negative employment effects for young women in Mexico due to PROGRESA represent a decrease from the baseline of about 10%, and positive effects for older women that are even larger (21% increase from the baseline for women older than 55), but can only statistically distinguish from zero the latter positive effect on the oldest group. These results are particularly interesting because their findings are consistent amongst themselves and with mine.

The studies considered in the review by Banerjee et al. (2017) include CCTs in Argentina, Brazil[3], Cambodia, Colombia, Honduras, Mexico, Nicaragua, Pakistan, and Philippines. Notably, children health and educational outcomes in all of these countries at the time re-

[2]These are all ITT effects, for PROGRESA (Mexico) they report the ATE effect, wich would represent a reduction of only 1.5%

[3]A standard CCT program relies on indirect measures of income, such that people's actual labor supply will not impact their chances of get the benefit. Argentina and Brazil however use administrative data to include current income considerations. The discussion in this paper applies to the standard CCT program, not to the type used by these two countries, that will create substantially different incentives.(Brazil: De Brauw et al., 2015; Ribas and Soares, 2011; Argentina: Garganta and Gasparini, 2015)

search was conducted were far below the same outcomes for these countries today (Bank, 2018). This accelerated progress may well be attributed to the CCT programs themselves. Effectively, CCTs have been proven very successful welfare programs on reducing poverty (Fiszbein and Schady, 2009), improve children's educational outcomes (Schultz, 2004; Maluccio and Flores, 2005), and access to health services (Gertler, 2004; Attanasio et al., 2005). However, this raises the concern that the beneficiaries are likely changing importantly over time. They have probably significantly higher educational attainment today than when the program was inaugurated, and their decision to work or not is most likely affected by this.

Continuing with the Mexican example, data from PROGRESA (currently Prospera) confirms that over time the beneficiaries of this early CCT program have greatly increased their schooling to a point in which they are more closely comparable to Chilean beneficiaries in 2007 than to their own nationals in the late nineties or early 2000s (their parents). According to Hernández Licona et al. (2019), household head schooling in moderately poor households receiving the conditional cash transfer has increased from 4.6 in 1994, three years before the program started, to 7.1 in 2016. Furthermore, for those in extreme poverty schooling increased from 3 to 5.8 years in the same time period. If the tendency to higher schooling of household heads remains, soon the beneficiaries of PROGRESA will be similarly educated than the beneficiaries of the Chilean SUF in 2007.

This is important because we can expect education to impact labor supply of mothers in several ways. Research shows that women that receive more schooling have less children (Schultz, 1993; Lam and Duryea, 1999), if they get at least eight years of education they become significantly more likely to work (Lam and Duryea, 1999), and invest more in their children's health and education (Strauss and Thomas, 1995). Additionally number of children, and the willingness to invest in their education without the assistance of a welfare program, will modify the relationship between the CCT and its beneficiaries in ways that could further impact their labor supply responses to it.

2.1.2 The Chilean Context

CCTs are effective because they motivate parents to modify their child investment decisions. They rely on simple rules to achieve their goal, and can do so because the desired outcome is easily monitored. However, over time the effect becomes less obvious. While the first generation might have had on average under five years of schooling, the second generation, brought up under the influence of the CCT, is going to be considerably more educated, and maybe even more so a third generation of beneficiaries.

As many CCTs are today over twenty years old, this becomes a pressing issue. For example, Mexico's net secondary enrollment has increased constantly since 1997, when its CCT PROGRESA was put in place. According to the World Bank, net secondary enrrollment in Mexico has improved from a little under 50% in 1996 to 77.7% by 2017. Further, the population over twenty five years old that has at least completed lower secondary, according to the same source, has increased from 33% in 1990 to 63% in 2018. A similar observation can be made for many other countries using CCT programs.

Chile's SUF was started in 1981. The World Bank data shows that by 1982 only 34% of the population over twenty five years old had at least completed lower secondary that year in this country. The same metric for the countries listed in Banerjee et al. (2017) was in the thirties by 1990 for all but the Philippines (46%) and Colombia (41%). Further, since 1997 these countries have exhibited a constant improvement of this metric: in Mexico, Brazil, and Colombia the population over twenty five that had at least completed lower secondary by 2018 was respectively 63%, 60%, and 54%; while the Philippines reached 70% by 2013. Meanwhile, Chile has been at 75% at least since 2004 (no information for earlier years).

The above means that parents today in countries that started CCT's in the 1990's are significantly more educated than their own parents. Therefore, they may not be in the same need to be motivated to provide their own children with education (Strauss and Thomas, 1995). In fact, the statistics on secondary net enrollment for these countries show that many of their children are going to high school. In 2017, 78% of Mexican children are enrolled, 86% of Chilean children, 90% of Argentinian youths, 79% of Colombian, and so on.

The changes experienced by these countries over the last two decades cannot be ignored by researchers, as they challenge the relevance of our work. Studies of these countries twenty or even ten years ago may be of questionable relevance today for a country that is so strikingly different from the country originally under study. In these circumstances, Chile can provide other countries with policy relevant information. This study, looking at Chile ten years ago, can provide useful information to one of the sixty countries using CCT programs today or in the near future.

This is an important consideration because we can expect the distortionary costs of a CCT program to vary over time, specifically as a function of school enrollment levels. This is so because the distortions that tend to increase labor supply depend critically on the preponderance of schooling among children. We may expect positive labor supply distortions to partially or completely offset negative distortions if schooling is low, and net distortions to be small or zero. However, if schooling is high it would tend to eliminate the positive labor supply distortions we recognize in theory, leading to larger negative net distortions. Therefore, distortions would tend to become more negative over time, countries that started a CCT program fifteen to twenty years ago would want to know whether their CCT program creates distortions in their current state, not in their initial situation.

2.2 The SUF program

The SUF program was established in 1981 with the explicit goal of increasing parents' investment in their children, and in 2015 reached about 15% of the country's households.

The program provides a monthly transfer per child and mother conditional on either health or educational investments. Children up to six have to be taken to regular medical controls. Older children have to attend school full time and be at most twenty four years old. If a family qualifies for the program, the per capita amount is the same for everyone.

These conditions are easily controlled for by the authorities, making the system efficient and relatively simple to enforce. However, the SUF also requires families to be part of the poorest 40 percent as measured by an index called *income generating capacity*. This index is built using self-reported information, making it very manipulable (Herrera et al., 2010; Irarrázabal et al., 2010).

From 2007 to 2010, this subsidy was made more generous and expanded its reach. The former resulted from an increase in the transfer's value. The amount transfered in 2007 was 31% larger than the previous year, and its subsequent growth was increased as well. From 2006 to 2015 the nominal value of the SUF grew at an average annual rate of 10%, 2.5 times faster than during the period 1998-2006 (Table 2.1). At the same time, it was turned into an entitlement, which lead to the program's doubling its reach by 2010 (Figure 2.1). The reason it takes some time for the CCT to grow is that the authorities need to estimate the income generating capacity of the beneficiaries in order to add them to the program, which means interviewing each family.

Overall, this overhaul to the SUF program could have important unintended effects as long as the program also meets two other requirements: being big with respect to some generally definable universe; and important enough to have noticeable effects on people's budget constraint.

Chile's SUF is very popular, reaching large sections of the country's families. It currently benefits over two million people in about 850 thousand different families, according to official data for 2016 (Subsecretaría de Seguridad Social, 2018). In a country with 18 million people, and 5.5 million households, this represents more than 10% of the population, and 15% of families. Furthermore, by decile of income, the SUF reached over 25% of the first two deciles and 17% of the third, according to survey data from 2015 (even with 30% undercount of SUF recipients).

On the second requirement, the SUF makes only a modest contribution to the budget of an average family. The transfer awarded to any particular *cause*[4] would represent only 4.1% of the national minimum wage in 2015, and even less in previous years. However, Table 2.2 shows that the benefits received through the SUF can represent a very substantive proportion of some people's work income.

This table shows three statistics, for two groups of SUF recipients. The first statistic shows the average value of the monetary transfer with respect to families' average work income; the second shows how many SUF recipients, with positive work income, receive between 20% and 100% of said income through the SUF; finally, the last column shows how likely is the SUF transfer to be larger than the recipient's declared work income, including those with zero work income.

I show these statistics for the general population and for young mothers between 18 and 24 years old, because I expect this group to consider the CCT transfer a more important source of income. What Table 2.2 shows is that there are some individuals that rely heavily on the SUF benefits. There is an important part of the SUF beneficiaries that would lose more than 20% of their family's income if the SUF is discontinued, ranging from 25% in general, to almost 50% for younger mothers.

There is one more important consideration to have with respect to the event that provides me with an identification strategy. At the same time the SUF doubled it beneficiaries, and more than tripled in total value, the entire welfare system was being expanded. The central

[4]The benefit defines causes and beneficiaries. Causes are children or mothers entitled to the transfer, and beneficiaries are the adults actually perceiving the benefit

government 'Subsidy' budget went from 4.8% of GDP in 2007 to 6.7% in 2010, and 7.6% for 2015 (Dirección de Presupuesto de Chile, 2017).

If other important subsidies saw very similar reforms as that of the SUF, then I would be identifying the response of the treated to this welfare system expansion, rather than specifically their response to the CCT program being expanded. Furthermore, many programs rely on the same *income generating capacity* index used to select eligible households for the SUF. However, a review of the most sizable subsidies that use this measure to determine elegibility can help mitigate this concern.

Most programs run by the government at this time are not a concern. This is because they either are addressed to a different population (pensions, scholarships) or because they are too small to be of significance nation-wide. However, a few programs are large and can target the same population that the conditional cash transfer program targets. The most salient example being that of housing subsidies. However, this subsidy can be received by both treatment and control, and running the model controlling for it does not alter my results, suggesting it is not an important issue.

However, there are other large programs that at the same time are addressed exclusively to mothers. The most likely to contaminate my results being food subsidies for children at school, which had a budget 65% larger than the SUF's in 2010. This subsidy even reaches a similar number of children as the SUF and was significantly expanded after 2006. However, this subsidy is given based on attendance to low income schools, and my geographical controls will allow me to control for the potential effect of this policy. Another program provides free childcare and preschool, and is about half the size of the SUF. However, this program only grew 11% in beneficiaries between 2007 and 2011 and would only benefit a fraction of the people that can benefit from the SUF.

2.3 Data

The data used come from seven editions of a survey created to evaluate public policy in Chile, called *Encuesta de Caracterización Socioeconómica Nacional* (CASEN, 'national socio-economical characterization survey'), covering the period from 1998 to 2015. This is a household survey first conducted in 1987, meant to be nationally representative of the population, that in 2015 reached almost 270 thousand people in 84 thousand households, and 100 thousand different nuclear families; effectively interviewing 1.5% of Chilean population.

The survey's purpose is to measure poverty, describe the poor, and guide and evaluate public policy, which is why it contains detailed personal demographic information, sources of income and labor force participation, among other things. Because it is meant to be repeated indefinitely there is a substantive effort in making versions comparable and the survey trustworthy. By 1998 it has been repeated five times before, giving people no specific reasons to believe the survey would be used against them if they admitted to improper behavior.

The analysis is based upon four surveys before the 2007 policy change: 1998, 2000, 2003, and 2006; and three surveys after the SUF's modification: 2011, 2013, and 2015. There is a 2009 survey that will be excluded from the principal analysis, because the expansion of the program is still incomplete at this point, and quality concerns[5].

Our relevant population are women between 18 and 60 years old, with at least one child younger than 24 (and full time student). This results in a sample of 319,298 observations for which some general statistics can be seen in Table 2.3. This table shows how the change in the SUF led to a sharp increase in the probability of having the SUF subsidy after 2006.

[5]The CASEN 2009 was excecuted by Universidad Alberto Hurtado, the four before and the three after by Universidad de Chile, This led to a sub-representation of 50% for SUF recipients in 2009, double the standard. Nevertheless, the results are robust to the inclusion of this survey, taking 2007 as the critical event year

Table 2.4 shows the same basic statistics for the subgroup that declares to receive SUF. These women are younger, less schooled, have more children, they have their first child sooner, are more likely to be single, work less and earn less than the overall population. However, their trends are those of the general population, and the table does not suggest a change in the composition of SUF beneficiaries. The biggest difference is as expected, a rather significant jump in schooling, likely a result of expanding the subsidy's reach.

It is important to analyze this issue a little deeper. If the selection to the program is significantly different after the reform then the analysis could be invalidated from the start. However, table 2.4 can help with this. First, it shows that the changes in the demographic characteristics of the people receiving the subsidy are smooth, which would suggest that the beneficiaries of the subsidy are not significantly different before and after the reform. Second, it also shows a significant jump in years of schooling for SUF beneficiaries. This effect can also be observed in figure 2.3, which shows that the probability of receiving the SUF increases more for more educated individuals. This observation maps perfectly to the selection mechanism for the subsidy, that rests heavily on declared schooling. Therefore, controlling for education among other observables should help control for this issue. Additionally, the survey data show that income considerations need not necessarily be an issue in this respect. Even after the reform, there is still households belonging to the first decile of income that are not receiving the subsidy (50%).

There is also one trend that seems to differ between the tables in a way that is not warranted, and that is labor force participation. Figure 2.2 shows how the labor force participation of the two groups (with SUF and without SUF) trend together only before the policy change. This figure also makes evident the smooth trend observed for non SUF recipients and should help alleviate concerns about group composition changes.The grey area in this graph is meant to signify the SUF expansion period, and the circles around the 2009 data points signify the alluded quality concerns.

The data also include detailed information about other outcomes related to labor supply that we will use to study the program's effects: labor force participation, average hours worked and average hours worked by the spouse (attributing zero to those without a job), average hours worked by those with positive hours, probability of working extra hours, and hourly wages. This six outcomes allow me to identify how individuals choose to respond to the change in the CCT program.

2.3.1 Definition of Treatment and Control

Ideally we could choose a random sample of mothers that benefit from the program to experience the reform, to use the unaffected mothers as controls. However, we do not have a group of beneficiaries that could not benefit from the program being expanded. This is particularly true because the program was expanded both in reach and value. This requires us to find another group that shares some similitudes to the group affected by the reform but that is not eligible for the subsidy.

In determining labor supply choices we may think some personal characteristics are particularly important: having or not children, educational attainment, income, age, being in a couple, are some examples. To choose control group we need to discriminate along the lines of at least one of these dimensions. Figure 2.3 shows that the probability of participating in the welfare program is a function of educational attainment for example. However, it also shows that for all education levels there is an increase in probability of receiving the subsidy.

Since a modification of the SUF could potentially affect any woman that has children eligible for the program. Nominally, this would be limited to mothers with eligible children in the first forty percentiles of per capita income, as the SUF should not be received by wealthier people. However, there is evidence that recipients misrepresent their situation in order to improve their chances.

Indeed, according to Herrera et al. (2010) the data used to determine eligibility to the SUF differs importantly to survey data for the same population. The head of household is older and more likely to be woman, and the family is smaller and much more likely to have someone with disability. Similarly, Irarrázabal et al. (2010) show in their table 4.2.1 that according to the data used to determine SUF eligibility, with two thirds of the country surveyed, 35% belong to the first decile of income. That would mean that at least 22% of Chilean population belong to the first income decile, which is only possible because the data is self-reported with the intent to improve chances of being eligible to the program.

Schooling cannot be used to discriminate mothers either. The survey data used shows that for all schooling levels mothers experience an important change in probability of receiving the subsidy after 2007 (Figure 2.3). The probability of benefiting from the SUF are very significant for any mother with 12 years of schooling or less (high school diploma), but are still relevant for 15 or 16 years of schooling, probably a consequence of a selection mechanism that relies on self reported and unchequed personal information.

Although we could define treatment and control groups by distinguishing mothers more or less exposed to the subsidy[6], there is actually a better option. There exists a group of women for whom the program is absolutely irrelevant, women with no eligible children. I will refer to them simply as 'not mothers', even though they could in principle have ineligible children (i.e. older tan eighteen and working, or older than twenty four). It turns out that this control group complies with the necessary conditions to be used successfully.

I will also be limiting my attention to the age range 18 to 50. The lower limit corresponds to the minimum age for mothers to benefit from the SUF, making this restriction a necessary one. The upper limit is made convenient due to characteristics of the data. The CASEN is a household survey that only started asking for total number of children after 20116.

[6]I actually did do this. In the appendix I define two alternative treatment/control definitions based one on the age of the youngest child and the other on number of children. The younger the children, the more children a mother has, the more she should value the transfer.

Therefore, limiting attention to women younger than 50 helps keep demographics between treatment and control more balanced, since older women typically no longer live with their young children. Nevertheless, the results are not overly affected by eliminating the latter age restriction (replacing it with 60, the retirement age for women in Chile).

Table 2.6 shows some general demographics for mothers and women without children. We can see here important differences between the groups: women without eligible children are on average younger, more educated, work more and are more likely to be single. These are all expected differences, but it makes salient the concern that these women may not be a good control for mothers.[7]

More importantly, figure 2.4 tests for parallel trends on labor force participation for both groups from 1998 to 2015. We can see here that the two groups, even though they have very different levels of labor force participation, have similar trends prior to 2007, which is the fundamental assumption that needs to hold for the difference-in-difference analysis I will describe later. Women without children are more likely to work, but do not seem to have reached a ceiling on labor force participation by 2015 either, which means there is no other reason for the series to separate after 2007. Furthermore, I test the parallel trend assumption for several subpopulations in figure 2.5 and do not find any violation of the assumption.

2.4 The Model

I use a difference-in-difference methodology to identify labor supply effects attributable to the SUF program being more generous. The preferred specification controls for several relevant covariates parsimoniously. Every regression includes controls for whether mothers have work experience, whether she has a spouse (de facto, regardless of legal standing), is part of the

[7]To address this concern I have alternative specifications for treatment based on how exposed mothers were to the CCT program. The results proved robust to these tests and are included in the Appendix.

primary family in the household (i.e. not the family of a son or daughter), and whether she is married; I also include dummies for year of survey (6), age and its square, family size, number of children, and fixed effects for neighborhood (359 dummies). If the regression is conditional on working I replace the dummy for work experience with a variable that records years in her current job. I also use heteroskedastic robust errors on estimation, given that differences-in-differences models are prone to underestimate them (Bertrand et al., 2004). Below is the model in its equation form.

$$y_{it} = \beta_0 + \beta_1 mom_i + \beta_2 \cdot post_t \cdot mom_i + \gamma_t + \Gamma X_{it} + \varepsilon_{it}$$

Where y_{it} is one of several outcomes of interest: labor force participation, weekly working hours, weekly working hours of the spouse (declared couple regardless of legal standing), log of weekly working hours, work overtime, defined as working over 50 hours a week (in Chile the working week was reduced from 48 to 45 hours in 2005), and hourly wage. $post_t$ identifies the timing of treament and is zero for data prior to 2007, and one for its complement. I also run this regression with three different interactions in order to evaluate possibly heterogeneous responses: whether the woman is single, whether she is younger (defined as in the age range 18 to 24), and whether she has a high school diploma (12 or more years of schooling).

As the results seem to be driven by all three characteristics, although with a clear distinction between the younger and the older groups, I run the regression combining two interactions. This allows me to closely identify the groups that are responding to the policy. It does not seem to be necessary to include more interactions in order to identify the groups reacting to the policy, but additionally it seems irresponsible to divide the sample in eight or sixteen groups for some of which the parallel trend assumption may not be valid anymore.

I study the trends of the series in figures 2.4 and 2.5. The first graph compares all mothers to all women without eligible children and shows that their trends are parallel prior to the reform. Figure 2.5 analyzes trends for several subpopulations: young women (figure 2.5a), older women (figure 2.5b), single women (figure 2.5c), women with a couple (figure 2.5d), the less educated (figure 2.5e), and those more educated (figure 2.5f). The parallel trend assumption holds for all these populations. However, women between 18 and 24 years old exhibit a very unstable pattern that goes up and down from one survey to the next that recommends caution when considering their results.

2.5 Results

I analyze the effect of the CCT reform on six different labor supply outcomes. Three of these outcomes are not conditional on working, with the first two utilizing the entire sample: labor force participation and hours worked attributing to those not working zeros; while the third, hours worked by spouses attributing zeros to non working spouses, only considers couples. The other three only apply to working women: log of working hours, probability of working overtime (defined as working more than 50 hours a week), and hourly wage. Each table presents these six outcomes in columns (1) through (6), in the same order as I just presented them.

Table 2.7 shows the effect of treatment on the population as a whole. This table suggest that the transfer created some distortions, as evidenced by a significant increase in spousal unconditional working hours and women's wages, but it is not clear exactly on whom. Taken together these two effects are not particularly sensible. A more natural response to higher wages for women within a couples is for women's labor to replace men's labor in the affected households. The effect on wages should lead to an increase in working hours by women, even those without a couple, yet we do not see a significant effect on the related outcomes. Furthermore, the effect on conditional hours and probability of working overtime are negative, even though not statistically significant. Nevertheless, this table confirms that the SUF is important enough to create distortions in the labor market. The fact that we cannot make sense of them from table 2.7 alone suggests heterogeneous responses.

Heterogeneous responses could make sense if for example the subsidy is only an important income source for some subpopulations, while not for others. We will explore this possibility by distinguishing mothers by age group. Table 2.2 shows that the youngest mothers, in the age range 18 to 24, are significantly more likely than the universe of beneficiaries to perceive the CCT transfer as a relevant source of income. Another reason why we could be observing heterogeneous responses is labor market rigidities and discontinuities. If women can only choose to work, for example, zero, twenty, or forty hours a week, then the response to treatment could be different for women that have a couple and women who do not. Single mothers, under these circumstances, may be unresponsive to treatment, if the transfer cannot offset a movement so large (forty to twenty hours, or twenty to zero hours). However, women who do have a couple may be more responsive, as their spouses could help offset the negative impact of their responses. More importantly even, while their own work is the primary source of income for single mothers, it is many times a secondary income source for women within couples, making the negative response much more likely for the latter group. We will consider yet a third possible source of heterogeneity, educational attainment. This will add the to circumstances stated above. A lower educational attainment being associated to lower wages, will more likely lead to larger labor supply responses to the transfer, as it becomes more or less valuable in terms of working hours.

Table 2.8 reports labor supply responses to treatment by age group, distinguishing mothers ages eighteen to twenty-four from mothers in the age range twenty-five to fifty. This table, and all others evaluating heterogeneity, directly reports effects by age group including the standard errors estimated using the delta method, and under these the baseline mean for each outcome. The most noticeable difference between these two groups is on labor force participation's response to treatment. Young mothers show a negative impact of 4.1% from their 2006 baseline, contrasted with no significant impact for older mothers. The only other statistically significant response comes from wages, that increase significantly only for the younger group. The rest of the outcomes all exhibit non significant effects. I find interesting that, other than unconditional hours (column 2) for mothers ages twenty five to fifty, most coefficients are far from being zero, even if they are all statistically not different from zero. This suggest to me that, as mentioned above, age is just one of the dimensions that is relevant for these effects, and there is probably still some important heterogeneity within each of these two groups.

Tables 2.9 and 2.10 share the same issues. According to the first of these tables, single mothers have no statistically significant response to the reform, and even though other point estimates are relatively large, they show significantly larger dispersion, which is possibly a result of internal heterogeneity. Meantime, the same table shows that women with couples have significantly lower internal dispersion resulting on three out of six significant responses, including an increase in labor force participation by almost 2%. The second table, by educational attainment, shows something similar. In this case, women that studied after finishing high school and work (columns 4-6) show more internal cohesiveness, with no significant response to treatment. However, less educated women that at most finished high school are internally very dispersed, if working, and show only a significant reduction in conditional working hours and an increase in wages.

These three tables, 2.8 to 2.10, prove that all these distinctions are important. In the first of these tables we see that the labor force participation response is different by age group. In the second table we observe that women in couples seem to be a relatively homogeneous group compared to single mothers, and these mothers in couples even increase labor force participation. Finally, table 2.10 suggest similarly that women that studied after finishing high school are mostly unaffected by the program's reform, but less educated women are affected, but do not form a homogeneous group. Overall, these three tables make it clear that we need to further subdivide these groups both to determine who is responding, and how.

The analysis until now shows that age is a critical determinant of labor force participation response to treatment. It also showed us that the effects may be concentrated mostly on the least educated population. In table 2.11 we look at the combination of these two elements, creating four different groups. Additionally, it showed that point estimates are larger on the intensive margin for single mothers, and larger on the extensive margin for those with a couple. This is looked further at in table 2.12 that combines age and single status, also creating four different subpopulations.

These two tables confirm that young mothers reduce their labor force participation in response to the program's reform. It further shows that this only happens if this mother is less educated (8% reduction) or single (4.1%). They also confirm a positive labor force participation reponse from older mothers, but only if highly educated or if they have a couple. These responses, although not as large as the negative ones, (1.2% for the highly educated and 2.2% for those in couples) are maybe more surprising, specially considering these are the groups less likely to be affected directly by the program (i.e. free their time from childcare or lose income from child labor). However, Cogan (1980) showed that married women (a subset of those with couples) can increase their labor supply in response to a similar treatment in the presence of fixed costs. This could be a response to an increase in wages if these women have non zero reservation hours (i.e. are not willing to work less than $x > 0$ hours a year). And, although we do not observe an increase in wages for the more educated women, we do observe one on those with a couple, a 4.5% increase. Further, the transfer itself could lead these women to lower their reservation hours, and in that way to become more likely to work.

The second outcome, unconditional working hours, is almost always statistically indistinguishable from zero. However, the only result that is significant is that of older mothers within couples. They, interestingly enough, lower their unconditional hours, although only by 1.7%. This is probably reflective of the responses of these mothers on the intensive margin. They lower significantly both their conditional hours, by 2.6%, and their probability of working overtime, by almost 11%. In 2006 the probability a mother with a couple worked overtime was 14.7%, and these mothers worked on average 30 hours a week more than the average for those not working over time. A 1.6 percentage points decrease in this probability translates to a loss of 0.48 hours per worker, that is a reduction of 1.1% in average conditional hours. This means that about 44% of the effect on conditional working hours is due to lower probability of working overtime, and the remaining 54% would be a reduction by individuals not working overtime, if we assume that other features did not change (which might not be the case, people working overtime may have reduced their hours more).

The third outcome on these tables, unconditional working hours of the spouse, shows no significant response when we subdivide the groups in this two ways. The effect is large for younger mothers, but very imprecise. The effect found for the overall group of mothers with a couple is still significant, and amounts to an increase of 2.5% in the spouse's hours. This effect could be a response from spouses of women that are now working, if they see their time as complementary.

The fourth outcome is working hours for workers, and we observe four groups that significantly reduce them, with the other four groups not showing a significant response. Critically, this outcome is harder to explain than the previous three, because it applies only to workers. This means that changes in the composition of workers can lead to increases or decreases in this metric, instead of all of the effect being attributable to the treatment itself. For example, the effect of treatment on wages for the four groups that respond reducing their working hours, if working, could originate on more productive workers that work on average less hours being attracted to the labor market, pushing the average down for this reason. Extra complexity stems from the fact that two of the four groups respond by shifts in their labor force participation to treatment. For these two groups compositional changes may be more important, and there is some evidence that compositional changes are actually important. For example, while wage increases almost the same for less educated young mothers and for less educated older mothers (8.4% and 8.7% respectively), the effect on conditional hours is about half for the latter group. This could stem from the negative labor force participation effect experienced by the younger group, if it pushes out of the labor market workers that on average work less hours. Nevertheless, likely some of the effect is not just compositional, as we observe negative responses for groups that increase labor force participation, sustain it, and decrease it. For all of the effect to be compositional it would be necessary for each group to be affected very differently by the same treatment, and in fact we can observe some consistent patterns. Decreases in conditional working hours seem more pronounced between the less educated workers, younger workers, and single workers.

In summary, we observe young mothers reducing their labor force participation and their conditional working hours, more so if they are less educated or single. We also see that older women increase their labor supply if they are in a couple or have more education, with the first group offsetting the extensive margin increase with reduction in conditional working hours and probability of working overtime.

2.6 Conclusions

This paper provides the first piece of evidence that CCT programs are likely creating labor supply distortions in labor markets today. I show here that we can expect both positive and negative externalities, depending on beneficiaries' personal characteristics as well as the relative size fo the transfer.

The analysis shows that younger mothers between 18 and 24 years old experience a relatively large reduction in labor supply that becomes more pronounced if they are less educated or single, on both the intensive and the extensive margin. This makes the distortion strongest for the people in more need, which is at the same time a logical conclusion and a concerning outcome.

We also observed an unexpected positive response by older mothers on the extensive margin, if they studied beyond high school and/or are part of a couple. However, these groups also seem to show negative intensive margin responses that are more in line with the prior expectations. These responses seem to be driven by the transfer itself and its effect on wages at least partially, and may warrant a look at the family unit as a whole to fully understand the dynamics at play.

These findings provide important information to countries that today have been using CCT programs for almost twenty years, and may be wondering about their implications today and in the future. Additionally, they stress the need for research beyond that done through randomized control trials to answer this type of questions, which can not properly be considered at the beginning of the process, but rather need to be asked much later and may require larger samples. Nevertheless my strategy has limitations, evident in particular in figure 2.5a for younger mothers (which shows a strange zig-zag form survey to survey), and makes more research of this type necessary in the future. Ideally, a look at a CCT program from a country that currently is poorer, and were the CCT maybe is relatively larger, could be very helpful to continue to understand their labor supply implications.

Progress needs to be made on studying other distortions as well. The incipient research on CCTs distortions has concentrated in labor supply, as this study itself. However, my results suggest that young mothers may be particularly affected by this type of welfare program. It would be interesting to see whether this demographic is also responding to the CCT by changing their education choices, fertility timing, and even family composition. On the latter choice, I noticed that my results differ if these controls are not included in the regression, which although probably is omitted variable bias, it could also indicate these choices are affected by treatment. Studying this further is necessary.

If these results are confirmed and expanded upon, CCTs could be redesign to avoid the most egregious distortions created. For example, my results suggests that increasing the age of eligibility for the program to 25 years old would, in the case of the SUF, possibly eliminate arguably the most concerning (labor supply) responses.

Table 2.1: SUF benefit per cause, by year

Year	**SUF**	**Growth**
1998	$3,025	-
2000	$3,310	4.60%
2003	$3,716	3.93%
2004	$3,797	2.18%
2005	$3,930	3.50%
2006	$4,126	4.99%
2007	$5,393	30.71%
2008	$5,765	6.90%
2009	$6,500	12.75%
2010	$6,776	4.25%
2011	$7,170	5.81%
2012	$7,744	8.01%
2013	$8,626	11.39%
2014	$9,242	7.14%
2015	$9,899	7.11%

Table 2.2: SUF benefits and Family Income

Year 2015	**SUF/Income**	**0.2<SUF/income<1**	**SUF≥income**
General	7.5%	14.5%	11.1%
Age 18-24	8.2%	19.6%	26.2%

The first column shows average value of SUF subsidy in terms of work income
The second column shows, for SUF recipients with positive income, what percentage of them gets over 20% of their income via the SUF
The third column shows, for all SUF recipients, what percentage gets more income through the SUF than through work

Table 2.3: Sample General Statistics

Year	Obs	Age	Years Educ.	# children	First Mom	% Single	% SUF	LF partic.	Work Income*	Work Hours
1998	35,095	36.9	10.0	2.07	24.1	21.4	11.2	44.7	658,369	17.0
2000	46,677	37.0	10.3	2.02	24.1	22.2	11.0	47.3	588,017	18.5
2003	47,703	37.6	10.6	1.98	24.3	23.6	11.4	50.5	544,443	17.7
2006	48,479	38.3	10.7	1.94	24.6	26.5	10.5	54.0	525,382	20.1
2009	43,506	38.9	11.0	1.90	24.8	29.4	13.3	54.5	512,965	19.2
2011	54,256	38.8	11.2	1.83	25.1	33.6	19.7	57.8	480,057	20.4
2013	39,985	38.9	11.5	1.79	25.3	35.6	20.5	60.0	425,274	21.4
2015	47,103	39.0	11.8	1.77	25.4	35.9	18.8	62.6	412,486	23.2

*refers to values for mothers working at least 1 weekly hour, in real value

Table 2.4: SUF recipients' General Statistics

Year	Obs	Age	Years Educ.	# children	First Mom	% Single	% SUF	LF partic.	Work Income*	Work Hours
1998	5,851	33.9	7.2	2.46	22.5	24.7	100	30.4	216,858	9.5
2000	10,223	34.1	7.5	2.31	22.7	27.6	100	32.0	183,378	10.1
2003	10,844	34.4	7.8	2.23	22.8	28.9	100	34.7	191,780	8.9
2006	9,885	35.0	8.1	2.17	23.2	31.4	100	39.8	197,561	11.6
2009	9,110	35.0	9.1	2.02	23.2	35.7	100	41.7	244,490	11.6
2011	12,708	35.0	9.5	1.95	23.5	40.3	100	46.9	209,973	13.4
2013	9,686	35.6	9.6	1.92	23.8	41.4	100	47.2	198,267	14.2
2015	10,745	35.6	9.9	1.95	23.7	40.0	100	48.7	183,421	14.7

*refers to values for mothers working at least 1 weekly hour, in real value

Table 2.5: At most High School Statistics (0 to 12 years of schooling)

Year	Obs	Age	Years Educ.	# children	First Mom	% Single	% SUF	LF partic.	Work Income*	Work Hours
1998	22,419	37.0	8.7	2.11	23.7	21.0	13.6	39.2	435091	14.8
2000	30,439	37.1	9.0	2.05	23.7	21.8	13.5	41.9	384101	16.1
2003	29,273	37.6	9.3	2.00	23.9	23.3	14.2	45.3	340891	15.4
2006	27,685	38.5	9.5	1.96	24.3	26.0	13.0	48.9	359295	18.0
2009	22,956	38.9	9.7	1.91	24.5	29.0	16.5	49.5	338108	17.0
2011	23,068	38.9	9.9	1.84	24.7	33.1	24.3	52.9	308800	18.0
2013	15,267	39.2	10.0	1.82	24.9	34.8	26.2	54.4	272577	19.3
2015	16,318	39.4	10.2	1.81	25.0	35.3	24.6	57.2	262733	20.6

*refers to values for mothers working at least 1 weekly hour, in real value

Table 2.6: Descriptive Statistics for treatment/control

Year	Obs	Age	Years Educ.	# children	First Mom	% Single	% SUF	LF partic.	Work Income*	Work Hours
Treatment										
1998	31,969	35.1	10.2	2.08	23.2	21.0	11.8	45.6	647,399	17.3
2006	43,159	36.3	10.9	1.96	23.7	26.5	11.2	54.7	508,901	20.2
2015	40,025	36.2	12.0	1.80	24.1	36.4	20.7	63.5	405,848	23.4
Control										
1998	3,495	34.4	11.0	-	-	35.4	0.7	62.8	843,445	25.2
2006	5,050	34.7	11.8	-	-	35.7	1.0	69.6	718,702	27.6
2015	7,109	34.5	13.1	-	-	35.5	1.4	75.8	534,704	29.2

*refers to values for mothers working at least 1 weekly hour, in real value

Table 2.7: Labor Supply Response Overall

	Unconditional			**Conditional**		
	Labor Force partic.	Hours worked	Hours worked others	Hours worked (logs)	Prob. overtime	Hourly wage
	(1)	(2)	(3)	(4)	(5)	(6)
treated	0.004	0.037	0.833^{+}	-0.012	-0.010	0.040^{+}
[s.e.]	[0.005]	[0.269]	[0.441]	[0.013]	[0.010]	[0.021]
mean	*54.7%*	*20.2*	*33.5*	*40.8%*	*13.1%*	*$2,517*
Observations	**308,993**		**220,756**		**126,404**	

significant at: ***0.1%, **1%, *5%, $^{+}$10%

All regressions include demographic controls. Unconditional outcomes control for whether the subject has work experience while conditional outcomes control for time in current job. The means used correspond to the year 2006.

Table 2.8: Labor Supply Response by Age group

	Unconditional			Conditional		
	Labor Force partic.	Hours worked	Hours worked others	Hours worked (logs)	Prob. overtime	Hourly wage
	(1)	(2)	(3)	(4)	(5)	(6)
treated						
18 to 24	-0.017**	0.256	3.344	-0.016	0.010	0.115**
[s.e.]	[0.007]	[0.451]	[16.88]	[0.024]	[0.014]	[0.036]
mean	*41.3%*	*13.2*	*23.4*	*40.1*	*11.6%*	*$1,385*
25 to 50	0.006	-0.047	0.159	-0.018	-0.016	0.037
[s.e.]	[0.005]	[0.307]	[0.467]	[0.014]	[0.011]	[0.023]
mean	*56.3%*	*21.1*	*34.7*	*40.8%*	*13.2%*	*$2,602*
Observations	**308,993**		**220,756**		**126,404**	

significant at: ***0.1%, **1%, *5%, $^{+}$10%

All regressions include demographic controls. Unconditional outcomes control for whether the subject has work experience while conditional outcomes control for time in current job. The means used correspond to the year 2006.

Table 2.9: Labor Supply Response by Single/Couple

	Unconditional			Conditional		
	Labor Force partic.	Hours worked	Hours worked others	Hours worked (logs)	Prob. overtime	Hourly wage
	(1)	(2)	(3)	(4)	(5)	(6)
treated						
single	-0.014	-0.176	-	-0.021	-0.018	0.050
[s.e.]	[0.009]	[0.508]	-	[0.020]	[0.018]	[0.032]
mean	*73%*	*27.6*	-	*40.1*	*12.3%*	*$2,048*
in couple	0.009*	-0.019	0.833^{+}	-0.010	-0.008	0.036*
[s.e.]	[0.004]	[0.181]	[0.441]	[0.011]	[0.006]	[0.015]
mean	*48%*	*17.6*	*33.5*	*42*	*14.6%*	*$2,774*
Observations	**308,993**		**220,756**		**126,404**	

significant at: ***0.1%, **1%, *5%, $^{+}$10%

All regressions include demographic controls. Unconditional outcomes control for whether the subject has work experience while conditional outcomes control for time in current job. The means used correspond to the year 2006.

Figure 2.1: **The SUF suddenly becomes more generous**

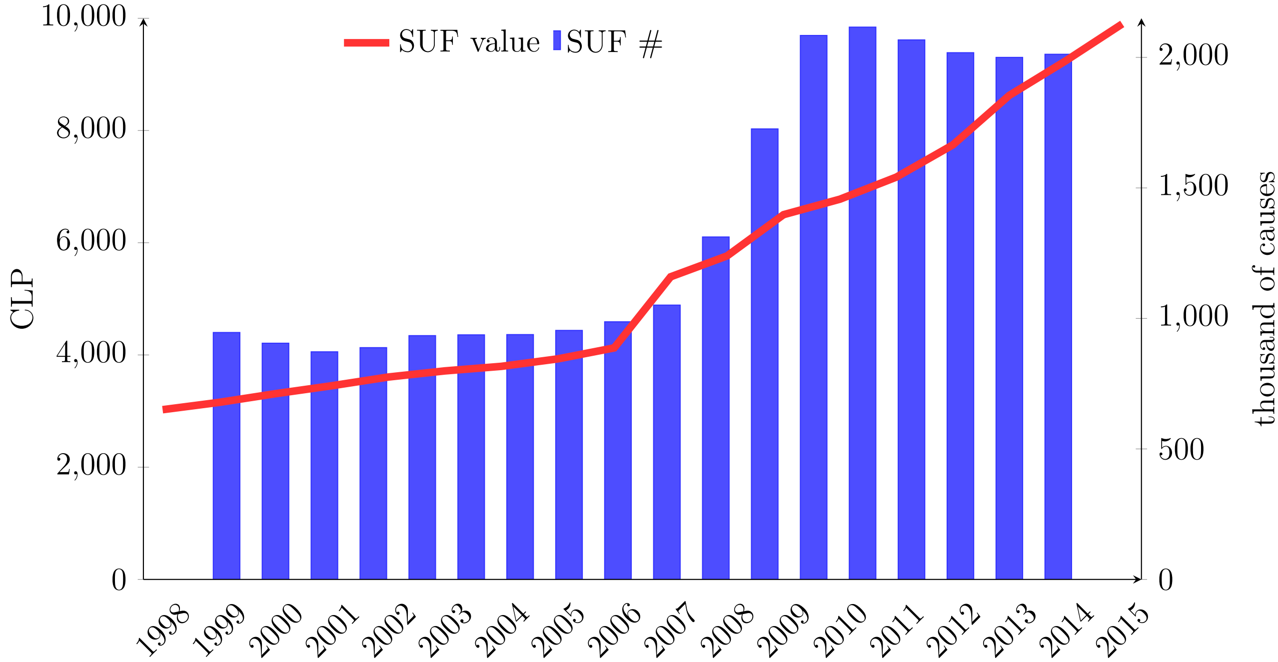


Figure 2.2: Mother's Labor Force Participation

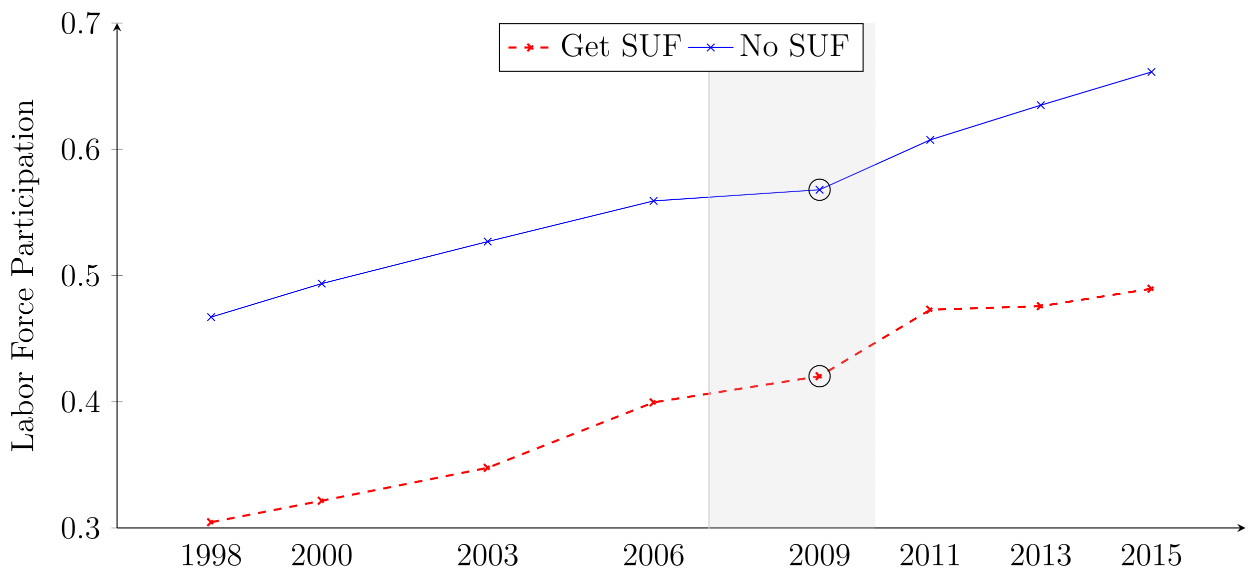


Figure 2.3: Probability of Receiving SUF by Schooling level

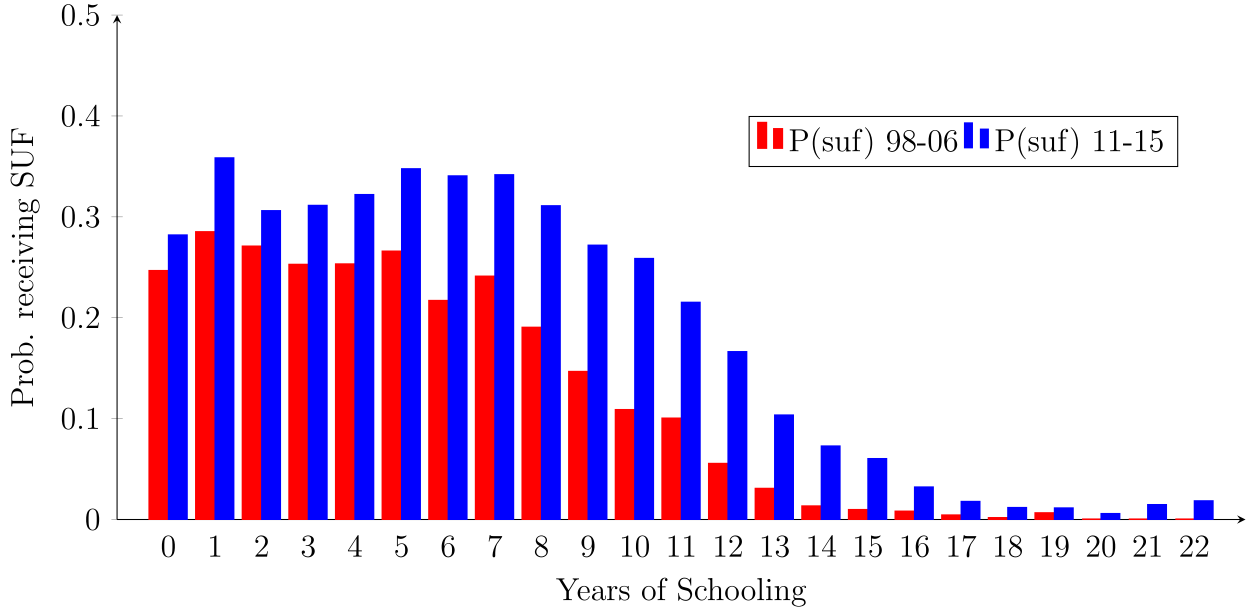


Figure 2.4: Test for Parallel Trends - Not Mothers v. Mothers

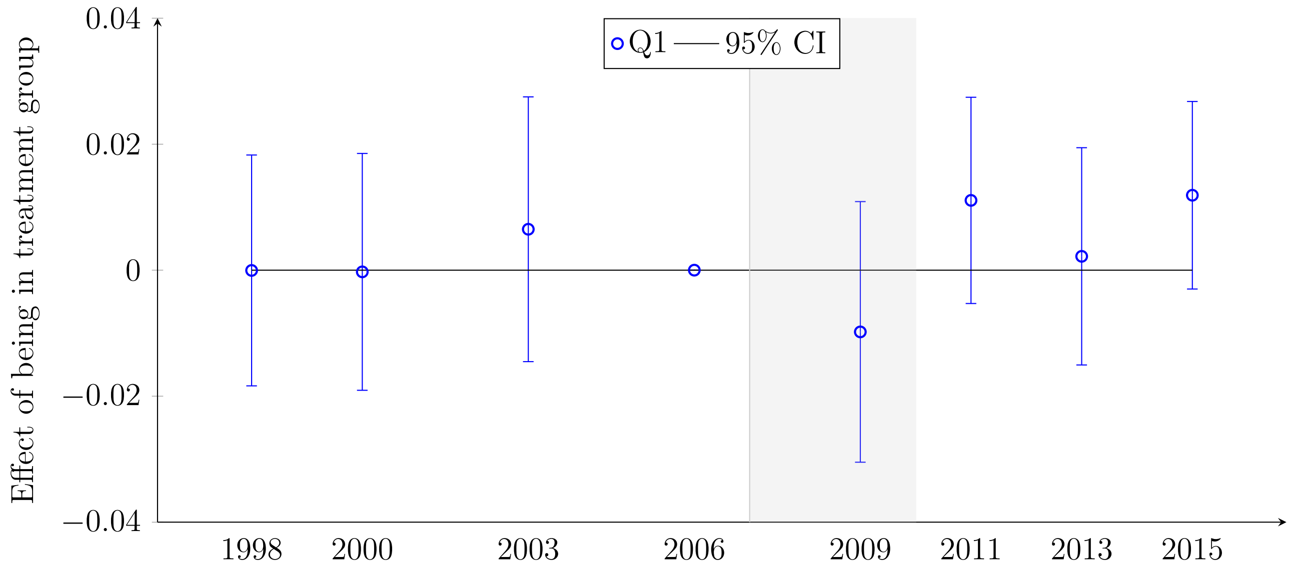
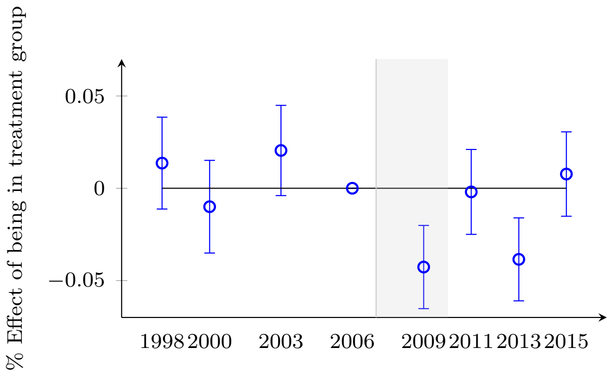
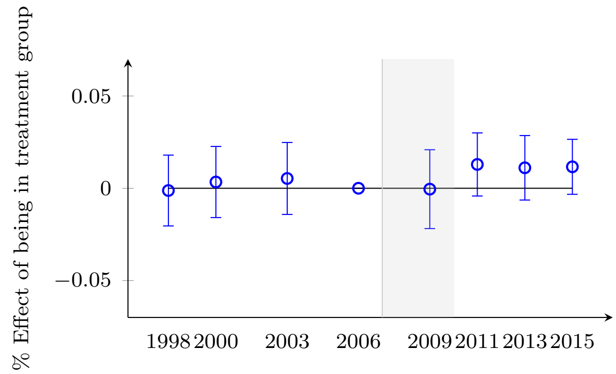


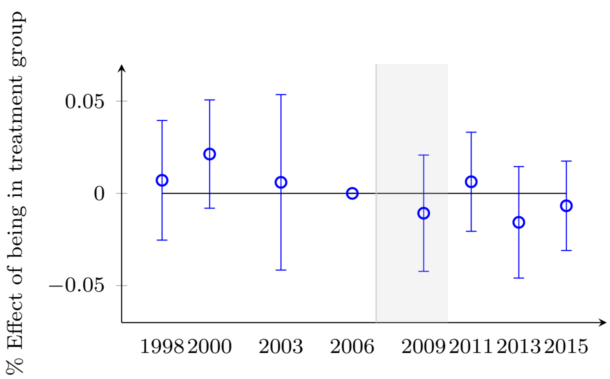
Figure 2.5: Tests for Parallel Trends for Subpopulations



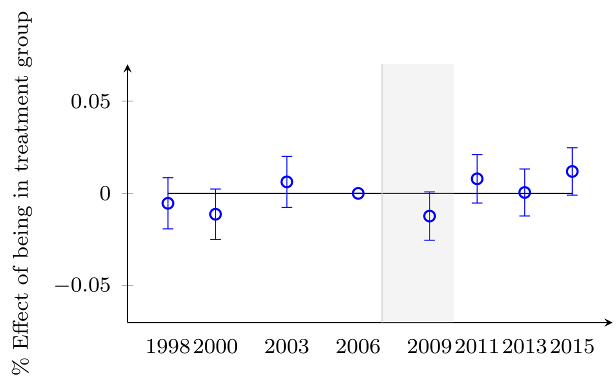
(a) Women 18 to 24 years old



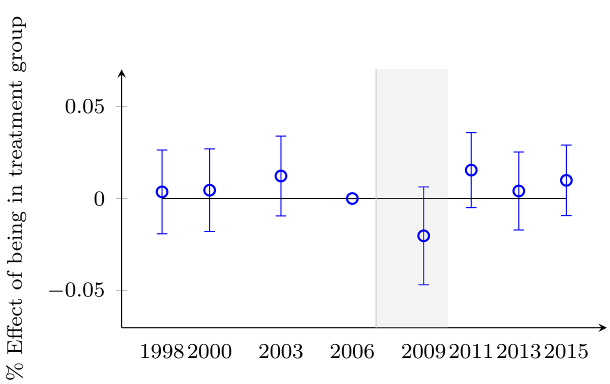
(b) Women 25 to 50 years old



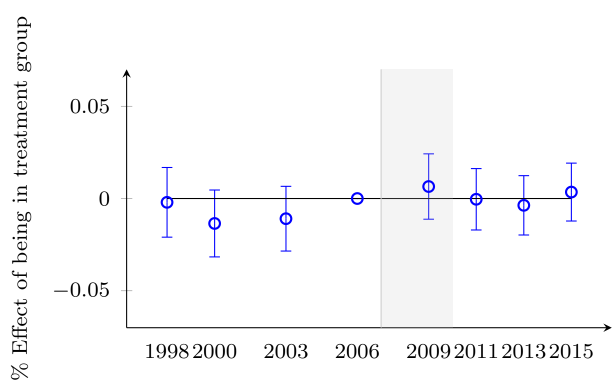
(c) Single Women



(d) Women in a couple



(e) Women with at most High school



(f) Women with more than High school

Table 2.10: Labor Supply Response by Education Attainment

	Unconditional			**Conditional**		
	Labor Force partic.	Hours worked	Hours worked others	Hours worked (logs)	Prob. overtime	Hourly wage
	(1)	(2)	(3)	(4)	(5)	(6)
treated						
≤High school	-0.0003	-0.212	0.397	-0.034^{+}	-0.011	0.080**
[s.e.]	[0.006]	[0.336]	[0.534]	[0.017]	[0.014]	[0.026]
mean	*49.7%*	*18.2*	*33.6*	*40.9*	*15.0%*	*$1,660*
>High school	0.006	0.398	0.922	0.014	0.001	-0.001
[s.e.]	[0.005]	[0.361]	[1.067]	[0.013]	[0.007]	[0.016]
mean	*72.1%*	*27.1*	*33.0*	*40.5*	*8.5%*	*$4,474*
Observations	**308,993**		**220,756**		**126,404**	

significant at: ***0.1%, **1%, *5%, $^{+}$10%

All regressions include demographic controls. Unconditional outcomes control for whether the subject has work experience while conditional outcomes control for time in current job. The means used correspond to the year 2006.

Table 2.11: Labor Supply Response by Age Range and Educational Attainment

	Unconditional			**Conditional**		
	Labor Force partic.	Hours worked	Hours worked others	Hours worked (logs)	Prob. overtime	Hourly wage
	(1)	(2)	(3)	(4)	(5)	(6)
treated						
18 to 24						
$\leq$ *High school*	-0.032***	-0.019	1.683	-0.068*	0.020	0.084^{+}
[s.e.]	[0.008]	[0.443]	[3.800]	[0.029]	[0.018]	[0.043]
mean	*40.5%*	*13.2*	*25.0*	*40.8*	*12.7%*	*$1,333*
$>$ *High school*	0.0001	0.539	7.447	0.070	0.012	0.126^{+}
[s.e.]	[0.012]	[1.042]	[2109]	[0.048]	[0.026]	[0.066]
mean	*46.0%*	*13.6*	*14.6*	*36.8*	*6.4%*	*$1,629*
25 to 50						
$\leq$ *High school*	0.005	-0.294	-0.042	-0.035^{+}	-0.020	0.087**
[s.e.]	[0.006]	[0.391]	[0.58]	[0.019]	[0.016]	[0.028]
mean	*50.9%*	*18.9*	*34.7*	*40.9*	*15.3%*	*$1,689*
$>$ *High school*	0.009^{+}	0.484	0.005	0.008	-0.001	-0.006
[s.e.]	[0.005]	[0.435]	[0.463]	[0.014]	[0.008]	[0.017]
mean	*74.1%*	*28.2*	*34.5*	*40.6*	*8.6%*	*$4,592*
Observations	**308,993**		**220,756**		**126,404**	

significant at: ***0.1%, **1%, *5%, $^{+}$10%

All regressions include demographic controls. Unconditional outcomes control for whether the subject has work experience while conditional outcomes control for time in current job. The means used correspond to the year 2006.

Table 2.12: Labor Supply Response by Age Range and Single/Couple

	Unconditional			**Conditional**		
	Labor Force partic.	Hours worked	Hours worked others	Hours worked (logs)	Prob. overtime	Hourly wage
	(1)	(2)	(3)	(4)	(5)	(6)
treated						
18 to 24						
single	-0.021^{+}	-0.465	-	-0.065^{+}	-0.002	0.255***
[s.e.]	[0.011]	[0.364]	-	[0.037]	[0.022]	[0.066]
mean	*50.8%*	*16.7*	-	*40.7*	*13.6%*	*$1,214*
in couple	-0.007	1.021	3.344	0.044	0.028	0.005
[s.e.]	[0.008]	[1.229]	[16.88]	[0.034]	[0.019]	[0.043]
mean	*33.1%*	*10.3*	*43.7*	*39*	*8.9%*	*$1,607*
25 to 50						
single	-0.012	-0.076	-	-0.016	-0.021	0.029
[s.e.]	[0.009]	[0.593]	-	[0.021]	[0.019]	[0.034]
mean	*78.5%*	*30.1*	-	*40.1*	*12.4%*	*$2,831*
in couple	0.011**	-0.304*	0.159	-0.026*	-0.016*	0.045**
[s.e.]	[0.004]	[0.152]	[0.467]	[0.012]	[0.007]	[0.016]
mean	*49.2%*	*18.2*	*45.7*	*42.2*	*14.7%*	*$2,152*
Observations	**308,993**		**220,756**	**126,404**		

significant at: ***0.1%, **1%, *5%, $^{+}$10%

All regressions include demographic controls. Unconditional outcomes control for whether the subject has work experience while conditional outcomes control for time in current job. The means used correspond to the year 2006.

Chapter 3

The principles' theory: Why some countries successfully increase their well-being while others remain poor.

We still do not have a thorough understanding of what determines the long term wellbeing of nations. The theory that institutions are the key determinant, though certainly an improvement over past theories based on geography or weather, fails to answer many important questions. For example, it seems to suggest all former colonies but four: United States, Canada, Australia, and New Zealand (UCAN) were given extractive institutions from the powers that colonized them[1]. At the same time, the per capita GDP by PPP ratios for former colonies in 1995 are 1:4.5:11:21, respectively for UCAN, America (not UCAN), Asia (not UCAN), and Africa. It is certainly true that the four rich former colonies are considerably

[1]Acemoglu et al. (2001) (AJR01) states 'This is in sharp contrast to the colonial experience in Latin America during the seventeenth and eighteenth centuries, and in Asia and Africa during the nineteenth and early twentieth centuries. The main objective of the Spanish and the Portuguese colonization was to obtain gold and other valuables from America. Soon after the conquest, the Spanish crown granted rights to land and labor (the *encomienda*) and set up a complex mercantilist system of monopolies and trade regulations to extract resources from the colonies' (p. 1375, no emphasis added)

richer than the rest, but it is also true that other American countries are considerably richer than Asian countries, that in turn were considerably richer in 1995 than African former colonies. The most critical shortcoming of the argument for institutions made by AJR01 is that they ultimately concentrate too much in the difference between four rich colonies (UCAN) and all other former colonies, to a point they became unable to see differences between the remaining ninety-eight countries. Case and point, if one subtracts the UCAN countries from their IV model the first-stage F statistic falls from twelve to under four (for log of population density, their stronger and more inclusive model) and log of population density becomes unable to predict average protection against expropriation risk (the coefficient falls from -0.21 to -0.08, with p-value of 0.46).

The theory that institutions are the key determinant of long term development also failed empirical testing. Acemoglu et al. (2002) (from now on AJR02) tested it by the use of an instrument, but the instrument itself was called into question by Albouy (2012). Although AJR02 respond to the challenge by Albouy (Acemoglu et al., 2012), they make clear in their response that the data for former colonies of Spain comes from sources investigating the XIX century, hardly a relevant statistic for XVI century Spaniards considering to settle or exploit America, or the Portuguese navigating around Africa to reach India. Regardless of whose arguments we find stronger, cross country analyses were temporarily abandoned in order to prove the theory in other ways[2]. However, I have found a way to evaluate the theory that is proposed in this paper taking advantage of data that is both more inclusive and less error-ridden[3].

I call it the Principles' theory. As the institutions theory, the definition of these principles is a complicated matter[4]. Principles should be understood as basic ideas that give rise to

[2]Recent research has focused instead on local experiments (Dell, 2010; Dell et al., 2018)

[3]Another important concern with the empirical analysis of AJR02 is that data limitations forces the loss of at least a third of the sample.

[4]AJR02 does not define institutions, instead they talk about *extractive institutions* and *institutions of private property*, which later they proxy by risk of expropriation.

all other beliefs we have. For example, we can believe private property is innate to human beings or that it is not, this would be a principle, that is a very basic idea that gives rise to all other ideas. At the same time the principal belief determines what we **can** believe about a lot of things but, at the same time, it does not create disjoint sets of beliefs. That is to say, the beliefs emanating from believing A and those from believing not A overlap, sometimes significantly. In general terms, these are ideas (principles) that are persistent over long periods of time and provide societies with ways forward, and tend to be seen as ends rather than means. Examples would be the nature of private property, the primacy of the individual/group, the (un)equal value of all persons, among others. The most important difference between principles and institutions is their complexity. Take Britain and Spain as an example, even today both countries defend private property, but while the former chose a common law system the latter uses a roman law approach. Consequently, if we are talking about institutions the British and Spanish differ in this respect, which is why it makes sense to treat them as different if we think institutionality is key; however, if we think the key to development are principles rather than institutions, we might consider Spanish and British as equal in this respect.

Principles will state a general direction to societies, but we need something else to understand contemporaneous differences between countries. A society's progress at any point in time depends not only on the direction taken (principles), but also the speed that is able to achieve in the pursue of its goals. Speed may vary even between societies sharing all the same principles, making it possible for us to observe different progress made at any point in time. I will show here that speeds indeed vary amongst former colonies headed the same way. I borrow the model by Alston et al. (2018) on state capacity build up to argue that some former colonies can reach higher speeds of progress due to the fact they were more cohesive at the starting point. This happens because within a society that is heterogeneous a section of the society can capture institutions for their own benefit and proceed to take advantage of the rest of society, reducing or even completely freezing progress in the process;

consequently, the less heterogeneous a society is the harder and less likely this is to occur. I show that the UCAN countries actually enjoyed this advantage, which helped them surpass other former colonies with similar principles and even their source countries eventually.

This theory can address the issue of large differences in current wellbeing between UCAN, America, Asia, and Africa, by providing a more complex and rich picture of the colonization process and how it differs for each continent and country even. I will show that the European made very different choices in their former colonies, which sometimes they stated explicitly. Furthermore, if I find enough evidence to support their claims, I generally accept their stated business in the new lands[5]. I further parametrize this new model to allow me to test it empirically against the theory of institutions as was presented by AJR02, and show that it can explain the data better.

My findings suggest that principles are a key determinant of long term development, and that the best seem to be those held by Europeans colonizers. One important implication of this theory is that the identity of the European -whether Spanish, French, British- is inconsequential for long term development. What matters is their intentions, either exploitative or constructive, and how these interact with the population currently inhabiting the land. Specifically, my model supposes that European principles are received when transmission is intended by the colonizer as evidenced in today's former colony, using to proxy transmission the date of founding of the first university. Newman et al. (1996) argues the purpose of the university is to keep and teach [all] our knowledge, a relatively popular view at least for these institutions in the past. However, the Europeans may integrate to native peoples or establish their own society, if the former is true then the society will have different groups within from the very beginning, but if the latter is true a society might enjoy cohesiveness[6] and progress faster; feature that I proxy by population density in 1500 for these former colonies. Taking

[5] I think it is a mistake to conclude that Spanish chose populated areas because they were looking for a work force (Engerman and Sokoloff, 2002), without even considering their stated reason to go there: the natives' conversion (Deus, 1537)

[6] Even an uncommon level of homogeneity if the society's elite remains by and large in their home country.

into account solely former colonies, countries with little or no population prior to the arrival of Europeans would be much more likely than other former colonies to have today societies erected by the new arrivals themselves, creating a homogeneous European society beyond the limits of Europe itself. The empirical exercise shows that if the transmission occurs, it explains a per capita GDP by PPP in former colonies 0.90 standard deviations higher, and if the former colony is cohesive the benefit is a per capita GDP by PPP 1.10 standard deviations higher.

To get to these results I combine two theoretical models, such that one helps determining direction and the other speed of progress. The first model, developed by Alston et al. (2018), allows me to determine a common direction among all societies that share the same principles. I purposefully use the term principles to differentiate from what Alston et al. (2018) call 'beliefs', which includes principles but also ideas that are not central to the system, and that can therefore differ between societies that share core values[7]. The second model, by Besley and Persson (2009), suggest that a few societies, originated by the colonization process itself, may have an advantage in terms of speed of convergence to their goal. Identifying causality will be possible because colonization will provide an exogenous shock to some former colonies' principles that is essentially random. However, we cannot use principles in an equation any more than we can use Institutions in one, so I created a proxy for them: year of formation of the first university. This is a particularly convenient measure of both the colonizer's intentions and of each society's exposition to European principles. Universities are meant to keep and teach our accumulated knowledge and so the implicit assumption is that universities flourish sooner in places were European principles were valued, kept, and taught (Newman et al., 1996). It is also available for all former colonies and much more objective and less

[7]The example given by Alston et al. (2016) of Brazil's progress from 1964 to 2016 divides this country's history in three, suggesting two cases in which 'Core beliefs' were affected. Brazil starts with Developmentalism (1964-1985), continues towards Social Inclusion (1985-1994), and finally arrives at Fiscally Sound Social Inclusion (1995-2016). However, at no point do the authors claim that the views on basic societal principles changed in any way. In fact, these three periods are treated as experiments to reach a goal that seems largely unaffected.

malleable than most indexes.

The rest of the paper is organized as follow. In Section II I will expand on the unanswered questions left by the institutions' theory and show how the principles' theory can answer them more satisfactorily. In Section III I will explain the theoretical model behind the principles' theory. In Section IV I present the historical evidence that some societies existing in former colonies before the arrival of Europeans suffered a shock to their own principles at the arrival of the Europeans, which led them to adopt their principles. In Section V I present the key results arrived at by the principles' theory, and show they exhibit a better fit than the original results from AJR02. In Section VI I compare my model to the model used by AJR02 to show the theory of principles can successfully challenge that of institutions. Finally, Section VII concludes and discusses ways forward.

3.1 Problems with the Theory of Institutions

There are many questions that the theory of institutions fails to answer. In this section I will cover the central concerns with this theory, and show how the theory of principles provides a better answer to them.

AJR02's take on colonization to support the Institutions' theory is excessively simplistic. Table 3.1 shows how three countries that were discovered by Europeans at roughly similar times, and endured some type of European domination after that date, exhibit nonetheless very different outcomes from said control. The American countries have cultural commonalities with their former European masters that are not observed in the case of India. The data on the ethnic origin of the population makes sense of this, as almost no Indian is in any way connected to Europe, but most Americans have either one or both parents from the old continent. However, there is also important differences between the United States and

Mexico. Namely, chances are that a Mexican chosen at random will have both American and European descent, but it would be very hard to find an American that has a comparable mix, most being only of European blood[8]. For better or worse, the Mexican is himself the integration of two peoples of very different background, while the American is instead the integration of different Europeans, with a common background.

Figure 3.1: European Colonization Differences

	US	**India**	**Mexico**
Colonized by	UK	Portugal/Netherlands/UK	Spain
National Language	English	Hindi	Spanish
Religion	Christian	Hindu	Christian
Population of European Descent only	72.4% (229M)	0%	10% (13M)
Population of Native Descent only	0.9% (1.1M)	97% (1.3B)	28% (35M)
Population of European & Native Descent	0.6% (0.9M)	0.06% (1M)	62% (78M)
Discovery	1565 AD	1498 AD	1517 AD
First University	1636	1857	1551
1995 per capita GDP (PPP)	$28,782	$1,485	$8,390
1500 Urbanization[1]	0	8.54	14.84

[1]Bairoch measure as used by Acemoglu et al. (2002)

I included in table 3.1 the value for the urbanization measure based on Bairoch used by AJR02. This measure shows that the 'reversal' advanced by these authors is true for the pairs US-India and US-Mexico, but it is not observed for India-Mexico, and although the reversal would be confirmed by population density for all three pairs, when available it is reasonable to take urbanization data as a better estimate of wealth/wellbeing. This is important because, while the correlation between GDP per capita and urbanization in 1500 is -0.39 for all available data (43 former colonies), it is only -0.2 (33 countries) between the countries with some level of urbanization, and turns positive for the 75% most urbanized (25 countries), at 0.13. Maybe the reversal is not a continuous response, but rather a binary treatment that allowed completely non-urbanized countries to surpass other countries.

[8]I am focusing here mostly on historical populations for these countries, somewhat ignoring the latest demographic turns for some of these countries that could not have impacted these countries' historical development.

This raises an important question for the theory of institutions. Since Mexico and India were colonized with extractive institutions (according to AJR02) maybe we would expect both to be similarly poor today, or maybe the reversal would not apply to these countries, as neither received good institutions. But if British institutions are better than Spanish institutions, how can it be possible for Mexico to be several times richer than India 500 years later?. Furthermore, Mexico was more tightly controlled by Spain than India by either of its colonizers, it was also controlled in its entirety, and for a longer time by some accounts. In fact, the measure of average protection against expropriation used by AJR02 gives a better score to India (8.27) than Mexico (7.5)[9]. This generates a complication for the theory of institutions, as the country with the 'better' institutions has a per capita GDP almost a sixth of the country with the worst institutions. In contrast, the theory of principles can make sense of this empirical fact, that countries in what was Spanish America are wealthier than former colonies in Asia or Africa.

Another fact that remains unexplained by the theory of institutions is how initial institutionality relates to future development paths. The central thesis in AJR02 is that the right institutions allow societies to progress. However, many times through history societies have thrown away their institutions to replace them with others, many times worst than the old ones. Therefore, even if some former colonies are given what these authors call *institutions of private property*, this alone should not be sufficient for all of them to achieve long term wellbeing. My point here is that institutions do not provide a society with a purpose (direction), it is the society (or its elite) that chooses its institutions according to its purpose, eventually discarding them if they fail to attain their desired goals[10]. Furthermore, the data shows that the countries with the 'right' institutions (UCAN) share with Europe much more than just institutionality, they used to be Europeans for the most part.

[9]It even gives a better average score for former colonies in Asia (7.0) than those in America (6.6), even considering Canada and United States.

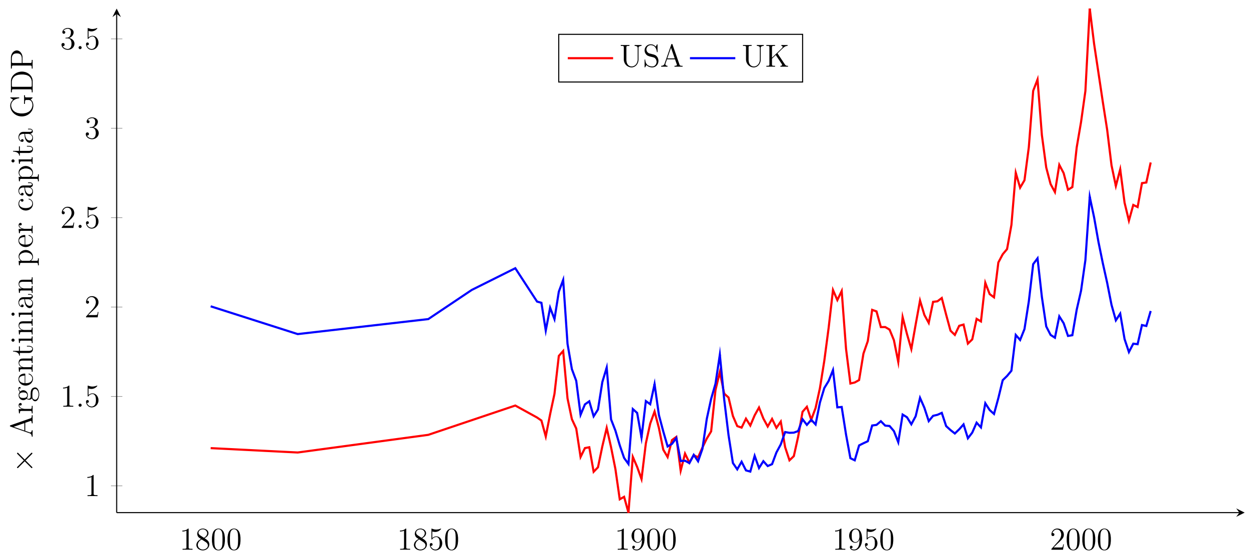
[10]This requires still more research from my part, but as shown by the Brazilian example used by Alston et al. (2016) in which Brazil changed its institutions three times in under eighty years, institutionality is not as constant as is typically assumed to be.

In fact, we should question the appropriateness of using the US-Mexico comparison over others available. If we want to measure the difference between receiving Spanish and British institutions then Mexico and US is a very uncomfortable pair. Population density in the 1500s was 0.3 standard deviations above former colonies' average for Mexico, and 1.9 standard deviations below the same average for the United States. A similar observation can be made with respect to urbanization, although in the reverse. Urbanization in the United States five centuries ago is estimated by AJR02 at 1.3 standard deviations below the average, and for Mexico is above the average by 1.7 standard deviations.

If we want to establish that Spanish and British institutions are significantly different in quality, Argentina is a much better choice than Mexico. Both Argentina and the United States have a zero on the Bairoch index on urbanization used by AJR02 and their population densities are also very close, with Argentina below the average for former colonies by 1.8 standard deviations (and the United States below by 1.9 standard deviations). Argentina, although richer than other former colonies of Spain in 1995, is still significantly poorer than the United States and Europe, but this was not always the case.

Figure 3.2 shows that in fact Argentina did not fall behind the United States until the advent of WWII, and remained in line with the UK until the 1970s (Maddison, 2013). If we refer back to AJR02's figure IVb, showing industrial production for a few countries, Argentina would likely be one more line jumping up with the UCAN countries, not one of the flat lines for Brazil, Mexico, or India. From 1800 until the advent of WWII the per capita GDP ratio between Argentina and United States hovered around 1.25 (average between 1900 and 1930 is 1.28) without any clear tendency to increase. To give context to this number, 2017 GDP ratio between USA and Germany was 1.34. Even more surprisingly, the GDP ratio with the UK remains at this level till the 1970s, less than half a century ago. This is important because AJR02 argue that the reversal was linked to the industrial revolution, which in turn is connected to protection of property rights. However, Argentina's development path is

Figure 3.2: Argentinian GDP per capita compared to USA and UK



problematic because it was given the same institutions than the rest of Spanish America and it thrived for much longer than AJR02 would have predicted. At the same time, it also fell like the rest of Spanish America, but not because of industrialization but much later for different reasons.

Critically, the theory of institutions failed empirical testing. Probably the stronger support to this theory came from the papers by AJR02 (2001,2002) . They managed to create an instrument for institutions (proxied by a measure of average expropriation risk) using data on settler mortality. However, we can identify at least two problems with their analysis. First, data quality issues, their database did not encompassed the entire sample of former colonies, and it imputed data for many countries. Albouy (2012) shows that original data existed only for 28 countries, with imputations for the other 36. He also showed that the imputations could be done differently, and that ignoring imputed mortality rates, and controlling for the type of mortality, makes the instrument ineffective. Furthermore, AJR02 make clear in their response to Albouy that data for American countries came from mortality estimates for at best early 1800s, full three hundred years after the Spaniards made the decision to exploit or build America. Additionally, an argument can be made squarely against the principle behind the settler mortality instrument, even under perfect quality data. If we believe that the Spaniards were at all sincere in their motif of converting the American natives, then they would have looked for populated places in the first place, and relatively ignore the empty ones (as indeed they did). If settler mortality is correlated with numbers of natives (more chances of wars and epidemics), perhaps a sensible assumption, it would make for a potentially confusing correlation between institutions and settler mortality.

AJR02 also use data on urbanization that has no information on Africa, covering less than half of their former colonies sample. It also has no Spanish or Portuguese sources for data on these countries' Empires. This creates a very real concern that the inclusion of missing data, and the finding of better sources, may alter their results significantly. The simple exercise of attributing to every African country an urbanization value of 0 (similar to the US, Brazil, Argentina, etc) would change the correlation between urbanization in 1500 and per capita GDP in 1995 from -0.4 to 0.18 (i.e. no reversal). Off course, we do not know how the data for urbanization for Africa would actually look, but the point remains that the 'reversal' seems to explain solely the movement of a few countries (UCAN) from the bottom to the top of the distribution, without shedding much light on the other 98 former colonies.

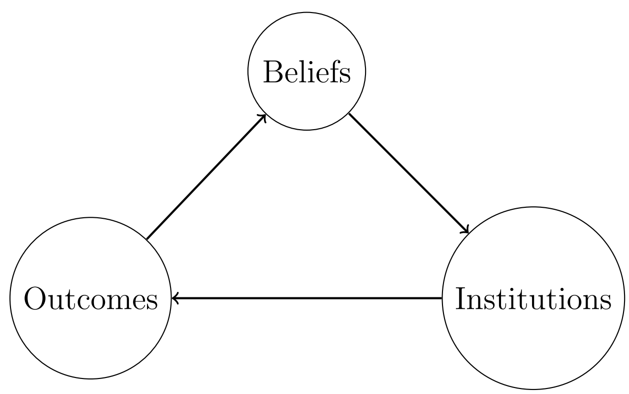
3.2 The Theory

The model I propose has two parts to it. The first addresses the way development is generally achieved by any society. The second concerns itself with the speed with which these societies can advance.

3.2.1 The Cycle of Long-term Progress

Lets consider the case for institutions further. Take figure 3.3, a simplification of the model by Alston et al. (2018) in which Beliefs, Institutions, and Outcomes determine each other in a cycle[11]. This simple diagram shows that an Institution that is developed by some society is intrinsically inseparable from the Outcomes of that same society over long periods of time (i.e. many such cycles). Which leads to at least two conclusions that are very problematic for the theory that institutions cause long term development. First, it questions the mere existence of a good instrument to institutions that can deal with endogeneity, given the very close relationship between Institutions and Outcomes. Second, it means that over time it would be increasingly hard for many independent societies to reach similar outcomes, as each previous decision affects the entire future of choices to be made.

Figure 3.3: Cycle Basic



Furthermore, it is very hard to sustain that a thing like institutions, which rise and fall over the course of human history, could provide enough stability by themselves to ensure the success of any society[12]. France killed her kings and the United States liberated her slaves,

[11]In their model Beliefs held by the Dominant Network (Government and other Organizations) determine Constitutional-level Institutions, which in turn determine Norms and other Institutions, which lead to Incremental change (from previous cycle), to economic and political outcomes. If the expected outcome is very different from the actual outcome, a window of opportunity can be open to reform Beliefs. They themselves illustrated this system with a set of reforms in Brazil in the 1990s that represented the belief in fiscally sound social inclusion.

[12]Understanding institutions as how a society acts upon and defends certain principles, which seems to

and both continued on their path of economic development and general success. Ignoring these past events, AJR02 suggest that similarities among the four successful former colonies (UCAN) have to do with institutional persistence (Albouy take up the same notion under the name of *autopilot*). My theory instead proposes that societies will correct their direction when they stray too far from their goals, even though it make take them some time and the correction is only realized when the deviation is relatively large, but this necessitates goals.

Societal goals are typically referred to as Ideals or Principles, and they are Beliefs. However, they are a small subset of beliefs that define all other beliefs[13], principles determine what can and should be done, and are much more stable than other beliefs, changing only over relatively long periods of time, and slowly. Since principles provide a goal and direction to a society, they must necessarily survive over many iterations of the cycle. Take the subject of private property, a notion in existence within human societies for literally thousands of years. As an ideal it could be considered a Principle, but how a society secures it and how far it believes private property should go, are malleable beliefs on which we and our parents may disagree even if we agree entirely on the principle of private property. Therefore, if we wanted to look at the same cycle in a 'short term' scenario (hundreds instead of thousands of years) we can depict it slightly different.

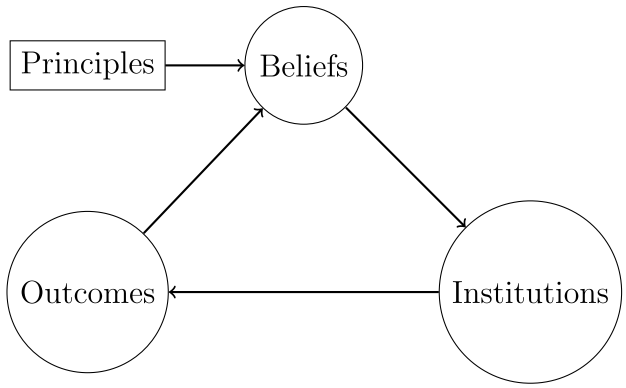
We can imagine that the cycle described by figure 3.3 is now, in the shorter term, informed by the society's principles (figure 3.4). In this case, if new principles are informed (or force upon) on a new society, they can restart the entire process. The new principles will redefine the set of acceptable beliefs, which in turn will lead to new Institutions, Outcomes,, and Beliefs, indefinitely. If we can identify a time in which new principles are assumed by a new society, an exogenous shock to that society, then we may be able to determine the benefit of

be how AJR02 understand institutions as well when they talk about 'institutions of private property' and 'extractive institutions', the principles being private property and extraction respectively.

[13] Alston et al. (2018) define core belief by their relation to Institutions: 'For us, beliefs embody a subjective view of how the beliefs inspire the formation of institutions with an expectation of how these institutions in turn will influence outcomes.' (page 280). In contrast, principles are higher, they are ideals that define our relationship to others, but do not pinpoint specific Institutions.

those principles compared to others. This is anything but a common development in history, but Europe's expansion starting in the XV century provides us with the necessary shock.

Figure 3.4: Cycle with Principles



3.2.2 The Principles

Before we talk about the other theory underpinning the principles' theory, we should discuss the notion of principles shortly. As I said previously, principles are very basic ideas (as axioms are basic) that make the basis for all other ideas we may have. Under this strict definition, a tax policy would not fall into the category of principle. Furthermore, notions that sometimes in economics are taken as basic are not principles either.

In this paper principles are understood to be general ideas (liberty, equality, private property). Their simplicity (as philosophers use the word) means that an infinite number of roads can remain relevant under their prescription, but also that they are never completely fulfilled or realized. Thus, not only do they determine which roads are feasible and which paths should not be taken, but also define societies' goals. By looking at people's most basic beliefs we are not going to be able to tell which tax rate their society will settle on for a transaction, but we will be able to tell whether it will (continue to) allow for purchases.

Under such definition, Wester European societies can be considered to be very similar between them from 1500 to today, even when their societies differ in important ways throughout the period. Furthermore, Nazis and Communist were not able to move most Germans and Russians away from their principles, and today the people of both countries believe still the same basic ideas shared by the rest of Europeans. The force of these principles is so strong that even Russians, for generations under communist rule, remain Orthodox Christians majority even though this religion opposes the most basic communist tenants (71% according to World Atlas)[14].

An important question that remains for principles is how are they assumed by a society. Can they be taken up in a continuum or should they be assumed binary: either you take them or you do not. Arguments can be made in both directions and it is hard to say which should be more convincing. People's beliefs are generally binary, if you believe one thing you deny others, but societies can be more complex and lead to a more nuance application of these beliefs. I chose to proxy principles in both ways, so as to be as transparent as possible. However, I personally prefer the binary understanding of principles because it additionally captures the intention to transmit principles from the colonizer.

3.2.3 Efficiency in the Cycle

Another important concern with respect to how countries evolve is how much each cycle allows it to progress. There is no reason to assume that even two countries with identical institutions and identical outcomes in any cycle will necessarily draw the same conclusions and update the cycle in the same fashion.

The model on state capacity building developed by Besley and Persson (2009) allows us to make sense of this fact and how it works. In this model the authors reach the conclusion that the ability to increase state capacity depends on three parameters: inclusivity of institutions,

[14] This is another point that necessitates more research from my part to strenghten it.

stability of power, and value of public goods. According to their theory there is two types of state capacity: the capacity to levy taxes, fiscal capacity; and the capacity to protect rights, legal capacity.

A society will invest in fiscal capacity if it is cohesive enough, and failing this, if it is stable enough. Cohesiveness has to do with how inclusive institutions are in this society. As a result, keeping all else constant, a mixed society would be less efficient in building fiscal capacity (or at best equally efficient). Further, a country that has no elite should be more efficient than another with one. If this is correct, then countries such as the United States and Canada, were Europeans started new societies in the XVII century, may have an advantage over all other countries for two reasons. First, because settlers can act under a guise of unique and almost complete cohesiveness when there is no other population being integrated to it. Second, because in the exercise of setting the institutions of these new places the elites can be given less advantages over other settlers (since everything has to be chosen again).

Is a little less clear that a more cohesive country will have the lead on legal capacity in this model, as other 'lower' types of states will also be very motivated to build fiscal capacity for other reasons. However, both capacities are considered complementary, an propositions 3.4 and 3.5 in Besley and Persson (2011), where the model is further studied, suggest that a country with more cohesive institutions would indeed build more legal capacity than another society in all other things its equal. If we believe that the integration of two cultures leads to a less cohesive institutionality (or at best a similar level), then this can become a difficulty for the resulting mixed society.

Summarizing, we are interested in two features of societies: their general direction towards development, and the speed with which they are able to advance towards this goal. Principles is what defines the former, the goals we set for ourselves provide us with meaning and direction. Cohesiveness explain how the latter may sometimes differ, societies that have less marked differences within will likely be able to acquire a greater speed.

This leads to the conclusion that even many societies that share common goals (Principles) can still have different ability to reach them faster, by virtue of them not dealing with sub-national groups that may manage to hijack the nation's institutions for their own wellbeing.

3.3 Evidence of Colonial Heterogeneity

'Other nations sent out bold explorers and established empires. But no other European people, before or since the conquest of America, plunged into such a struggle for justice as developed among Spaniards shortly after the discovery of America and persisted throughout the sixteenth century' (p.1 Hanke, 1949)

In this section I will answer a few questions that are important to establish that indeed there are different types of colonies (some which should not even receive that name) and that ignoring this likely will lead to important error. First we look at whether there exist some example of early colonizers that did not merely attempt to exploit natives of foreign lands, but rather integrated to them. Next I show that these early 'colonizers' were at odds with the exploitation theory in both words, written and spoken, and acts. Finally, I argue that this reality completely destroys the usefulness of settler death rates as an instrument.

In 1492 Christopher Columbus discovered a new continent ignored previously, but never again, by Europe and the World. It was a very momentous time for Spain that would see her reach the top of the world. The discovery of America would mark the beginning of intense discussion in Spain on what to do about this newfound territory and what right had Spain to possess it, if any.

Insua (2018) summarizes the philosophical struggle that followed the discovery. According to Insua, America's discovery lead to an important theological effort to determine whether Spain had a right to stay in America, and at what capacity. Agustinism being replaced by

Tomism, it is no longer accepted by theologians that pagan kingdoms have no legitimacy (questioning therefore the validity of the Alexandrine Bulls that gave the right of conquest to Spain). Additionally, the natives of the new lands were not enemies of the Christendom (as were the Muslims expelled from the peninsula the same year of 1492), which means there was no just reason to do war against them. This theological discussion did not only run in the background as conquistadors kept doing their business without being bothered, but expressed itself in tangible acts during the XVI century and the whole of Spanish tenure in America. The object of the conquest itself was considered to be the wellbeing of the natives, that before were involved in acts that the Spanish considered immoral (human sacrifices and cannibalism among these). The often and important disruptions to the 'encomienda' system by which the Spanish attempted to manage the natives, and the cornerstone of their American economy, provides us with a measure of the importance of this states purpose (native's wellbeing).

In the initial period of conquest that Zavala (1935) calls 'Antillian', before 1542, the 'encomienda' institution suffered many changes and interruptions. It was outlawed in 1502 by Nicolás de Ovando, only to come back the following year after the effort was declared a failure[15]. In 1511 the threat by Dominican priest Montesinos to deny absolution from their sins to 'encomenderos' (de Las Casas) leads the king to organize a board of jurists and theologians, and the release of the Laws of Burgos by December 1512 that imposed stronger regulation to the 'encomienda' system. This was followed by the 'requerimiento', a both tragic and comical document that introduced a totally useless formality in which the indian was given the chance to accept the Spanish rule in order to avoid war, a sample of the Spaniard's extreme legalism. Nevertheless, its existence is illustrative of how important it was for the Spanish king to protect his conscience from what even then were deem abuses of the natives. This is followed by many experiments that attempt to find an alternative

[15]'Because of the liberty the indians enjoy they run away from the christian and do not work.' (author's translation, Royal Order December 20, 1503)

method to the ‘encomienda’, recorded in Hanke (1949), that nevertheless fail to produce the expected results.

Furthermore, as soon as 1526 the power of the conquistadors was being seriously limited. They were to travel with two ecclesiastics approved by the Council of the Indies and could not enter into war against natives without their consent. These ecclesiastics would also make sure the conquest was just in all other matters, including paying for native’s property taken by Spaniards (Hanke, 1949). The King even stopped the conquest in 1550 with the express objective of making sure the process would be carried on justly, with the proper respect of natives and their possessions. On this respect the author comments ‘Probably never before or since has a mighty emperor in the full tide of his power ordered his conquests to cease unitl it could be decided whether they were just.’ (p. 117 Hanke, 1949).

Estrugo (2002) makes the point that Spaniards, unlike British (‘Saxons’) mixed with locals both in America and Africa (page 20). Indeed, both Pérez (1947) and Gregorio Marañón in his prologue to Pérez de Barradas (1976) recognize this difference, Pérez by distinguishing the portuguese ‘position’ colonization from the Spanish ‘rooting’ colonization; and both by arguing that this supposes the conservation and integration to the local population that is so common in Hispanic America but not present in North America or Brazil. Campa (1931) describes the effort put into learning their languages, in order to advance the Spanish cause but also to convert them. Errington (2008) not only stress the importance of missionaries in linguistic projects, but specifically the importance and scope of the work done during the XVI century in Mexico and Philippines, two to three centuries prior to similar contributions from other nations (protestant missionaries). Further, Elliott (2007) counts 225 towns and cities founded by the Spanish in America by 1580 with a Spanish population of 150,000. Kagan et al. (2000) and Bayle (1952) too refer to the importance of founding cities and municipalities in the Spanish America. Lummis et al. (1989) makes the point that Spain not only discovered America but also civilized it. According to Lummis et al. Spain had hundreds of cities all over America (including San Agustin, FL and El Paso, TX) while France had only timidly run some unfruitful expeditions, Portugal sustained some small towns 'of little importance', and Britain had done absolutely nothing (page 81).

Direct evidence that the Spanish in America were not looking exclusively to exploit the peoples and lands they discovered can also be found in the recopilation of all laws promulgated for the Indies and Philippines before 1681 (Consejo de las Indias, 1943). In the first volume we can find laws that attempt to secure for the natives access to hospitals, schools and universities, the catholic faith, justice, and even good graves. For example, a law from 1630 prohibited ecclesiastic judges to decide on cases of natural law involving non-catholics. Another, prohibited judges to give pecuniary punishments to natives 'Because of their extreme poverty, of which we wish to free them from' (author's translation, p. 81 Consejo de las Indias, 1943). The law that creates the universities of Mexico and Lima states specifically its purpose to educate natives 'it is convenient that our vassals, subjects and natives may have in them Universities and general Studies in which they may be instructed and graduated in all sciences and faculties...' (author's translation, p. 191 Consejo de las Indias, 1943), and there was also a law that created schools specifically for native chieftains' children.

Off course, this is not to say that Spaniards behaved like saints towards the native American. There is broad recognition that conquistadors abused their power just as much as the rest of the European colonizers, and not always followed the written law. However, there is also evidence that they did integrate to the peoples they found in the places they discovered, unlike many European colonizers [16].

[16] The strongest evidence being the fact alluded before that large fractions of the current population in former Spanish colonies are a mixture of native Americans and Spanish settlers

3.4 Empirics

To test if the transmission of European principles and the cohesiveness of societies can explain long term economical development it is necessary to find a good of measure of both characteristics. In the previous section I showed that there is current and historical evidence that some Europeans did not limit themselves to extract from their foreign lands, but rather chose to populate and build them. This latter type of colonizer would not impose extractive institutions, but rather its own institutions, the institutions that they themselves thought best. Furthermore, they would transmit something even more important to them, their own Principles, their Ideals and long term goals, to this integrated society of natives and Europeans.

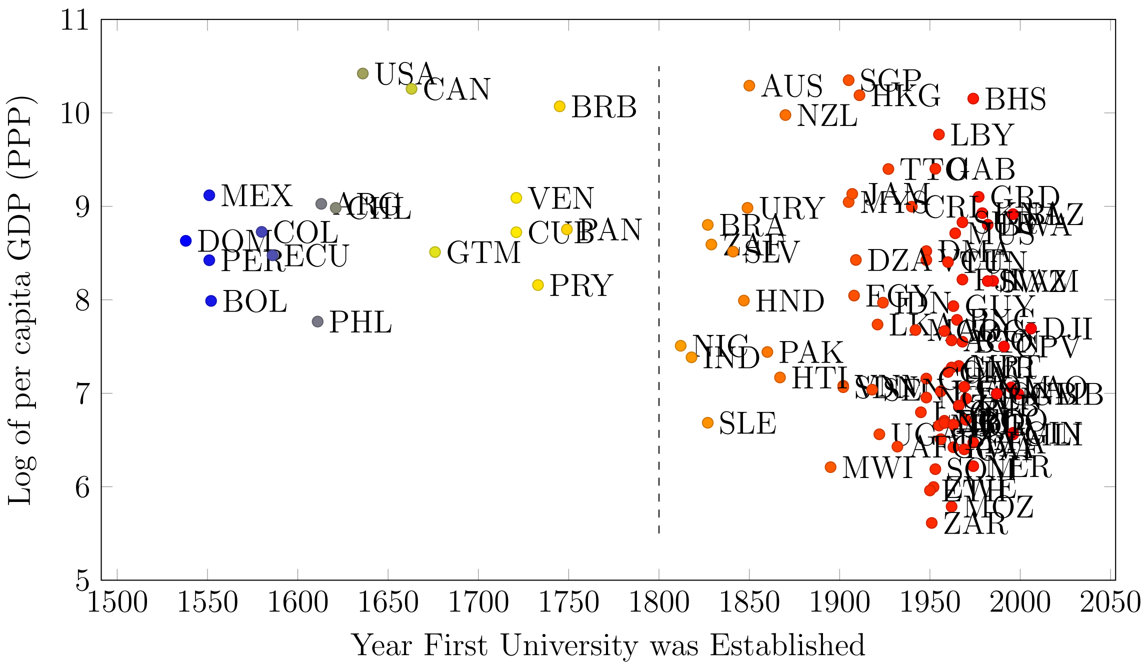
I have found a convenient proxy to measure this transmission, with relative advantages when compared to most historical measures: it results in no missing data for the sample of former colonies; and can be easily checked today, and even improved upon in the future. I will create two different measures for principles based on this proxy, one binary and one continuous. This is because although I would sustain a binary proxy is probably more appropriate, there are good arguments to sustain either is correct.

3.4.1 Principles by Education

Educating natives using European institutions is a particularly transparent way to transmit principles. We can use the dates at which universities were established for the first time in each colony to create a measure of the length of time that the population was influenced by European ideals. We will allow for a caveat however, if the population is of majority European descent (not mixed, solely Europeans) kthe variable assumes the value one, as transmission only makes sense on non-European peoples. The implicit assumption here is that if a former colony has a population that is of majority European descent, this means that the society was in fact established by these Europeans themselves, rather than them integrating to a society with a different set of principles that was already in existence at the time of their arrival.

Figure 3.5 shows in the x-axis the year the first university was founded in each country, and in the y-axis its 1995 per capita GDP by PPP. We observe clearly that countries that had a university sooner are likely to be middle or high income countries in 1995, never poor. A strong correlation is present, -0.41 for the entire sample and -0.36 without the four that conform the group 'neo-Europe'. Overall, suggestive evidence that European principles matter for long term development.

Figure 3.5: Correlation between year of founding of First University and per capita GDP by PPP



I find that one important concern with this measure is that I do not have detailed information as to who was actually able to attend these institutions always. It is entirely possible that universities were closed to Natives for a long time, meant for Europeans only. However, this is less of a problem given that most of the oldest universities belong to former Spanish colonies that were legally open to natives (Consejo de las Indias, 1943)[17].

Another potential concern has to do with country divisions that sometimes are determined after the establishment of universities. This is particularly concerning for small states that previously were part of a big Empire. For example, although Spain established more than twenty universities in Spanish America, neither was founded in modern day Paraguay or Puerto Rico. It is perfectly reasonable that the Spaniards thought it unnecessary at the time, and not really indication that they treated those two geographical locations differently. However, instead of making assumptions about this historical fact I choose to leave the variable as it is. Further, Paraguay still gets its first university relatively early, the year 1733, which makes the argument for manually altering the measure even less defensible.

Using date of formation of the first university I created a binary variable that takes the value one when European principles are sustained, as proxied by the early founding of a university, and zero when the first university appears too late (after 1800). Additionally, when the population is predominantly of European descent the binary value takes value one[18], as transmission in this case is not necessary.

In case principles are rather assumed on a continuum, making the binary variable a relatively bad measure of their presence on a society, I also created a continuous measure of principles based on the year of founding of the first university:

$$Principles_c = \frac{1995-Founding_c}{1995-First_univ}$$

On this formula, country's c Principles is a ratio between how long the university has been in place with respect to the colony in which the first university was established. It will take

[17]Nevertheless, this would be an interesting subject to explore further in the future to improve on this measure of exposition to European principles. My next efforts in this direction will focus on determine that the differences in law did in fact led to different practices.

[18]The score -0.99 is that of New Zealand, the most densely populated of the UCAN countries in 1500 according to the data

the value one for the former colony that received the first university the year 1538, and zero if the university was established in 1995. A negative value means that the university was established after 1995, which occurs in five occasions. However, for countries were the population is predominantly of only European descent (for example UCAN), transmission makes no sense and the variable's value is replaced by one.

There is one additional complexity that I did not address here, and that is the starting point of colonies. Colonies were started in 1500 in some places, but not until 1800 in some others, maybe this should enter somehow in our principles' proxy. My view on this issue is that in actuality all colonies could have started around 1500 or 1600. The Portuguese and the Spanish went around the world before the end of the XVI century and chose not to go into Sub-Saharan Africa, and a little later the same can be said of British, French, and Dutch. Since I understand the lack of African colonies early on as a choice, I prefer to leave the index as it is, for both binary and continuous proxies of principles.

3.4.2 Cohesiveness

Theoretically a society without elite, everything else the same, has an edge over other societies because it is more cohesive. In theory this would allow it to advance faster, as there is a lower risk that a subgroup will use the state for their own benefit instead of advancing the general good.

The uniqueness of the colonial times is that it seems to have created exactly this situation in a few countries. If Europeans were attracted to other lands were native population was very scarce, then the result could be a very cohesive society. In principle, these societies can be even more cohesive than their source. The new society can redefine everyone's role in it, take power from their home elites under the right circumstances, and pick and choose what features of the old world to take and which to leave. This could also go the opposite direction, and the new society be even less cohesive, if the elites manage to take for themselves even more privileges. If the former is true, we should see that my proxy for cohesiveness has a positive impact on long-term development, if the latter is true we might see no effect or even a negative one.

To take this into account my model includes a 'Cohesive' variable. This variable takes the value one if population density in 1500 was particularly low[19]. In particular, considering the group christened by AJR02 as 'neo-Europe' I chose a cutoff that keeps this four countries in this group. If log of population density in 1500 is less than -0.99 then the variable takes the value one, zero otherwise. The results are not overly sensitive to the choice of cutoff. This is because there are not many countries close to the cutoff. Only seven countries are considered non-cohesive while being relatively close to the cutoff, and of them only 2 are closer than 20 percentage points to the cutoff. However, even including these seven countries as cohesive the measure remains significant, while moving the cutoff to the left makes the coefficient even larger.

Table 3.1: Former colonies by log of population density (rounded)

log(population density)	frequency
-4	2
-2	10
-1	8
0	38
1	20
2	10
3	9
5	1
Total	98

AJR02 use the continuous version of this variable, the actual data on population density for each country. I will explain with two examples why I think the proper measure should be

[19]Engerman and Sokoloff (2002) recognize this idea in page 4 'The logic is that great equality or homogeneity among the population led, over time, to more democratic political institutions, to more investment in public goods and infrastructure, and to institutions that offered relatively broad access to economic opportunities,'. Off course, Besley and Persson (2009) not only recognize the same idea, but model it too.

binary. My binary index based on population density is meant to only flag countries were Europeans could set up their own society ignoring the natives, so that the new society is European in every respect but geography. If this is not the case, then they will have to integrate or choose to exploit the locals; in this case it does not matter if there is one or one hundred million natives, because the society will not be cohesive in either case.

3.5 Results

Following AJR02 I will present tables of results with six different sets of controls, which I took from them directly: no controls, weather controls, by resources, by colonizer, by religion, and by continent. Additionally, the main tables (tables 3.2 and 3.3) have four and three panels respectively. Panel A of both tables test the relevance of principles by themselves, then I put to the test the importance of cohesiveness in panel B of table 3.2, finally I put both halves of the model together in panel C of table 3.2 and panel B of table 3.3. In the last panel of each table I test whether the date of founding of the first university by itself (without my considerations for transmission) gives similar results[20].

Panel A and B from table 3.2 (and panel A from table 3.3) show that both halves of the model behave well. Their goodness of fit is comparable to AJR02's (column 0) with or without controls[21] , and the coefficients are strongly significant and additionally stable across specifications. Even the relatively larger estimates for principles in panel A of table 3.2 compared to those on table 3.3 can be attributed to the fact that the continuous measure will have a lower average than the binary one. To note is the fact that the variable 'cohesive', a simplification of population density in 1500, seems to achieve a slightly better fit than the variable from which it is derived, and its standard errors show no lose of precision when

[20]That is, without changing the value of the 'principles' variables to 1 when the population in a country is majority of European descent

[21]There is only a small difference between the results published by AJR02 due to some small corrections (i.e., I changed the Philippines' colonizer identity to Spain) and the updated GDP data

compared to log of population density.

However, panel C of table 3.2 and panel B of table 3.3 show that the model works significantly better complete. The coefficients for both principles and cohesiveness are reduced somewhat, but remain strongly significant and very relevant in terms of their size. Furthermore, the fit is improved significantly with both variables in the regression. The adjusted R^2 increases from 0.22 to 0.34 in the specification without controls, and is significantly higher for all the specifications with controls also, except for one (controls by continents).

With respect to column (6) in both tables, with controls by continents, it should be noted that this is a potentially confounding set of controls given the history of the colonization process. The process started with Portugal navigating around Africa (mostly ignoring the continent in terms of conquest) and Spain conquering almost all of America during the XVI century with a similar idea about what to do there; additionally, later African colonization in the XIX century was executed with similar motivations by the Europeans involved. Indeed, the first university in Africa was founded in 1827 while most American countries of significant size had a university by 1800 (with the exception of Brazil that got its first university on 1827).

Panel D in table 3.2 and C in table 3.3 provide an additional robustness check by simplifying the definition for the proxies for principles. The variable 'University by 1800' will differ from the binary definition of principles when the first university comes after 1800 and the country is majority of European descent (Australia, Brazil, New Zealand and Uruguay are now considered to not have European principles). The change impacts more former colonies in the continuous definition, as all countries with population that is today majority of European descent were moved to principles value of 1, some from original values over 0.7 (Argentina, Canada, United States), others from values closer to 30% (Australia, Brazil, New Zealand, and Uruguay). However, the results remain very significant and stable, even though we do observe the coefficient is reduced for the binary definition and the standard errors are larger for both models (with no controls it is 15% less precise under the binary specification, and 25% less precise under the continuous one).

My preferred specification is found on panel C of table 3.2. I prefer this specification because the intention to keep/transmit European principles to the settlers and natives was clear on many early European colonizers and because I tend to think that principles cannot be taken lightly or partially. If this is the case, the binary definition used in my preferred specification is more logical and sensible. However, it also requires me to choose a cutoff for the variable, potentially adding bias into the model. For this reason I tested the sensibility of my results to changing the chosen cutoff for founding of the first university. My results show that moving the cutoff by 50 years in either direction changes little, the model remains highly significant (only in one specification significance is lost for principles) and stable, even at 1700 and 1900 cutoffs the model remains explicative, although the coefficients tend down and standard errors up, the latter is particularly true for the cutoff of 1700.

The key take away is that these numbers confirm the theory laid down in this paper. Having European principles and being cohesive seems related to better long term outcomes for former colonies. A former colony that has the right principles will have a per capita GDP in 1995 that is between 0.74 and 0.92 standard deviations higher, and a cohesive one will have a per capita GDP in 1995 that is between 0.98 and 1.10 standard deviations higher. A country that meet both conditions will have a per capita GDP that is between 1.72 and 2.02 standard deviations higher, a very significant impact. Lets take the example of India, which did not received the principles nor is it cohesive by this model's data. Its 1995 per capita GDP by PPP was $1,611, but it would be $4,697 if it had received the British (or Dutch, or Portuguese) principles, and it would be $17,410 if the country was additionally cohesive (an impossible outcome however). By the same token, if Mexico was cohesive in addition to having European principles its per capita GDP by PPP in 1995 would not have been $9,123 but $33,801 that is it would have a per capita income representing 100% of the 1995 GDP

Table 3.2: Model with Controls - Principles (binary)

				Controlling by:			
Panel A	AJR02	None	Weather	Resources	European	Religion	Continent
	(0)	(1)	(2)	(3)	(4)	(5)	(6)
Pop. dens.	-0.39***						
in 1500(log)	(0.07)						
Principles		1.41***	1.19***	1.24***	1.46***	1.61***	0.44
(binary)		(0.26)	(0.28)	(0.25)	(0.34)	(0.32)	(0.29)
F	27.96	29.64	4.95	11.24	6.20	10.21	18.12
adj-R^2	0.22	0.22	0.39	0.49	0.30	0.29	0.41
				Controlling by:			
Panel B		None	Weather	Resources	European	Religion	Continent
		(1)	(2)	(3)	(4)	(5)	(6)
Cohesive		1.72***	1.43***	1.46***	1.60***	1.70***	1.21***
		(0.31)	(0.30)	(0.28)	(0.30)	(0.32)	(0.27)
F		30.18	5.37	11.70	7.85	11.31	26.12
adj-R^2		0.23	0.42	0.50	0.36	0.31	0.50
				Controlling by:			
Panel C		None	Weather	Resources	European	Religion	Continent
		(1)	(2)	(3)	(4)	(5)	(6)
Principles		1.07***	0.93***	1.01***	0.88***	1.10***	0.20
(binary)		(0.25)	(0.27)	(0.23)	(0.35)	(0.33)	(0.27)
Cohesive		1.31***	1.17***	1.22***	1.26***	1.25***	1.17***
		(0.30)	(0.29)	(0.26)	(0.32)	(0.33)	(0.27)
F		26.88	6.45	14.63	8.11	12.34	20.91
adj-R^2		0.34	0.49	0.58	0.40	0.38	0.50
AJR02 adj-R^2		0.22	0.40	0.46	0.33	0.21	0.50
				Controlling by:			
Panel D		None	Weather	Resources	European	Religion	Continent
		(1)	(2)	(3)	(4)	(5)	(6)
Univ. by		1.20***	1.00***	1.06***	1.05**	1.26***	0.23
1800		(0.30)	(0.33)	(0.28)	(0.41)	(0.38)	(0.30)
F		16.32	4.07	9.32	4.27	6.29	17.34
adj-R^2		0.13	0.33	0.44	0.21	0.19	0.40
Obs	96	100	99	98	99	93	100

significant at: *10%, **5%, ***1%.

Using data from St. Louis FRED on per capita GDP by PPP to replicate AJR02 results. AJR02 adj-R^2 was estimated by me for log of population density in 1500 using updated and corrected data.

Table 3.3: Model with Controls - Principles (continuous)

				Controlling by:		
Panel A	None	Weather	Resources	European	Religion	Continent
	(1)	(2)	(3)	(4)	(5)	(6)
Principles	1.77***	1.47***	1.59***	1.70***	2.04***	0.54
(continuous)	(0.33)	(0.35)	(0.33)	(0.43)	(0.39)	(0.36)
F	29.86	4.88	10.97	5.79	10.76	18.08
adj-R^2	0.23	0.39	0.48	0.28	0.30	0.41

				Controlling by:		
Panel B	None	Weather	Resources	European	Religion	Continent
	(1)	(2)	(3)	(4)	(5)	(6)
Principles	1.29***	1.06***	1.26***	0.85*	1.35***	0.13
(binary)	(0.33)	(0.35)	(0.31)	(0.45)	(0.42)	(0.34)
Cohesive	1.25***	1.13***	1.20***	1.29***	1.18***	1.18***
	(0.31)	(0.31)	(0.27)	(0.34)	(0.34)	(0.28)
F	25.14	6.11	14.04	7.56	12.06	20.73
adj-R^2	0.33	0.47	0.57	0.38	0.38	0.50
AJR02 adj-R^2	0.22	0.40	0.46	0.33	0.21	0.50

Panel C	None	Weather	Resources	European	Religion	Continent
	(1)	(2)	(3)	(4)	(5)	(6)
Univ.	1.72***	1.48***	1.49***	1.53**	2.07***	0.24
exposure	(0.41)	(0.45)	(0.39)	(0.61)	(0.53)	(0.42)
F	18.00	4.23	9.35	4.22	7.49	17.22
adj-R^2	0.15	0.35	0.44	0.21	0.22	0.40
Obs	100	99	98	99	93	100

significant at: *10%, **5%, ***1%.

Using data from St. Louis FRED on per capita GDP by PPP to replicate AJR02 results. AJR02 adj-R^2 was estimated by me for log of population density in 1500 using updated and corrected data.

per capita of the United States instead of only 27%.

3.6 Principles v. Institutions

In this section I show further evidence that my model can challenge the theory of institutions successfully (as presented by AJR02). To do this I use the same variables with which AJR02 argued their 'reversal' to represent their theory, and the two variables I created to represent the theory of principles.

The first two tables on this section seek to directly compare the two theories. Each table has six outcome columns, the first three include the theory of institutions proxied by population density in 1500 and the last three include this theory proxied by urbanization in 1500. Table 3.6 uses the same data used by AJR02, and table **??** uses more inclusive data obtained from the Saint Louis Federal Reserve Bank.

Table 3.6 uses original data from AJR02, with the only difference coming from a few corrections to the data that do not even change the point coefficient. Columns (1) and (4) show their results, (2) and (5) include the principles model with the binary proxy, and (3) and (6) include the principles model using the continuous measure. For the more inclusive data used, population density in 1500, we see that including the principles' model takes away more than half of the predictive power of population density, but remains significant at 5% level. However, for urbanization rates, the variable is turned completely irrelevant when we add the principles' theory, even switching signs. Another noteworthy feature is that the inclusion of the principles' model leads in both cases to significantly better fit, as measured by the adjusted R^2.

The institutions model does not deal well with the inclusion of this new model. With the updated data on per capita GDP by PPP log fo population density and urbanization in 1500 have no significant effect on long term wellbeing as measured by GDP in 1995. Tables 3.4 and 3.5 use data from the St. Louis FRED on GDP by PPP and allow me to recover 5 more colonies for population density and 2 for urbanization. Table 3.4 shows that adding either variable for 'principles' reduces the coefficient on population density 30%. Including instead the 'cohesive' variable the coefficient is reduced by 45%. Furthermore, when the entire model is included, the coefficient in population density in 1500 loses statistical significance[22]. A similar dynamic is obtained from table 3.5, although the data quality makes both models suffer. Even though less than half of the former colonies remain, principles manage to remain significant in all but one case (column 6). The variable 'cohesive' shows much better behavior, but 'urbanization in 1500' even shifts signs when the entire principles model is included. Overall, the 'principles' variable seem to have a harder time with this variable than with population density in 1500. However, I attribute this to the sample reduction, which is clearly not random. As I mentioned before, there is no data on Africa nor there are any Spanish or Portuguese sources for data on their Empires in Bairoch, the source for these estimates.

One might be concerned that 'cohesive' may interact badly with 'population density in 1500' because it is derived from it. To test whether the addition of cohesive was creating technical problems for population density in 1500 I tried a few exercises. First I gave the same treatment to continuous principles, dummies that took value one for only its tails, I tried left and right tails and neither affected the significance of principles. Then I added a dummy that is one for the right tail of the distribution of population density in 1500, but it also had no effect on the coefficient for population density (which became -0.40 significant at 1%). I conclude that the reason is probably because the extra information contained in population density in 1500 is not informative to the econometric model, and under the principle of parsimony the dummy should be preferred over the continuous measure.

[22]If instead of using 'principles' I use 'University by 1800', the results are similarly significant, but the coefficient on population density only falls by 15%

Table 3.4: AJR02's model (pop. density) & Principles' model together - Updated Data

X	(1)	(2)	(3)	(4)	(5)	(6)
		Log of GDP per capita by PPP 1995				
Pop. dens. in 1500	-0.39***	-0.27***	-0.27***	-0.21**	-0.09	-0.13
	(0.06)	(0.08)	(0.07)	(0.10)	(0.10)	(0.10)
Principles (binary)		1.02***			1.02***	
		(0.28)			(0.27)	
Principles (continuous)			1.32***			1.24***
			(0.34)			(0.26)
Cohesive				1.07**	1.07**	0.89**
				(0.45)	(0.42)	(0.32)
F	27.96	22.66	23.44	17.53	27.18	26.68
adj-R^2	0.22	0.31	0.32	0.26	0.47	0.46
Obs	96	96	96	96	96	96

significant at: *10%, **5%, ***1%.

Using data from St. Louis FRED on per capita GDP by PPP to replicate AJR02 results.

Table 3.5: AJR02's model (urbanization) & Principles' model together - Updated Data

X	(1)	(2)	(3)	(4)	(5)	(6)
		Log of GDP per capita by PPP 1995				
Urbanization in 1500	-0.08**	-0.06*	-0.06*	0.0002	0.02	0.01
	(0.06)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Principles (binary)		0.55*			0.48*	
		(0.29)			(0.25)	
Principles (continuous)			0.76*			0.51
			(0.38)			(0.34)
Cohesive				1.40***	1.35***	1.29***
				(0.36)	(0.35)	(0.36)
F	7.22	5.63	5.89	12.55	10.17	9.38
adj-R^2	0.13	0.18	0.19	0.36	0.40	0.37
Obs	43	43	43	43	43	43

significant at: *10%, **5%, ***1%.

Using data from St. Louis FRED on per capita GDP by PPP to replicate AJR02 results.

Table 3.6: AJR02's model & Principles' model together - Original Data

X	Log of GDP per capita by PPP 1995 - source: AJR					
	(1)	(2)	(3)	(4)	(5)	(6)
Pop. dens. in 1500	-0.38***	-0.16**	-0.18**			
	(0.06)	(0.08)	(0.07)			
Urbanization in 1500				-0.08***	0.02	0.01
				(0.03)	(0.03)	(0.03)
Principles (binary)		0.87***			0.64***	
		(0.20)			(0.21)	
Principles (continuous)			1.06***			0.68**
			(0.26)			(0.29)
Cohesive		0.73**	0.62*		1.31***	1.22***
		(0.32)	(0.32)		(0.29)	(0.30)
F	46.12	27.18	26.68	9.35	15.85	13.37
adj-R^2	0.33	0.47	0.46	0.17	0.53	0.48
Obs	91	91	90	41	41	41

significant at: *10%, **5%, ***1%.

Tables 3.7 and 3.8 include both models at the same time together with the five different sets of controls. These regressions confirm the prior results completely, population density is never significantly impacting per capita GDP in 1995 by PPP but cohesiveness is significant always, with a stable impact on long term development, and principles is only not relevant for the continent controls, the same set of controls that confused the model in the previous section.

I decided to also replicate table VIII from AJR02 in which they present their instrumental variables results. Although my contention is that their instrument is not going to be able to deal with the endogeneity problem, because of how they are intertwined with outcomes as well as data quality issues, we can still test the strength of the principles' theory in this scenario. However, principles and institutions are also related, so we would not expect the instrument to be innocuous to the model.

Indeed, we see that principles and even cohesiveness suffer from the inclusion to the IV model. Principles are significant on either stage in five of the six regressions, but in two

Table 3.7: Model with Controls - Institutions & Principles (binary)

	Controlling by:				
	Weather	Resources	European	Religion	Continent
	(1)	(2)	(3)	(4)	(5)
Pop. dens.	-0.06	-0.02	-0.11	0.03	-0.04
in 1500	(0.11)	(0.09)	(0.10)	(0.12)	(0.09)
Principles	0.90***	1.00***	0.82**	1.13***	0.10
(binary)	(0.28)	(0.25)	(0.36)	(0.33)	(0.28)
Cohesive	1.03**	1.18***	0.96**	1.37**	0.99**
	(0.45)	(0.37)	(0.43)	(0.47)	(0.38)
F	6.03	12.62	7.30	10.34	18.38
adj-R^2	0.49	0.58	0.40	0.39	0.52
Obs	95	94	95	89	96

significant at: *10%, **5%, ***1%.

Using data from St. Louis FRED on per capita GDP by PPP to replicate AJR02 results.

Table 3.8: Model with Controls - Institutions & Principles (continuous)

	Controlling by:				
	Weather	Resources	European	Religion	Continent
	(1)	(2)	(3)	(4)	(5)
Pop. dens.	-0.09	-0.06	-0.14	-0.01	-0.04
in 1500	(0.11)	(0.09)	(0.10)	(0.12)	(0.09)
Principles	1.05***	1.254***	0.85*	1.41***	0.002
(continuous)	(0.35)	(0.32)	(0.46)	(0.43)	(0.36)
Cohesive	0.88*	1.04***	0.88*	1.17**	0.98**
	(0.46)	(0.37)	(0.45)	(0.48)	(0.39)
F	5.83	12.26	6.99	10.22	18.33
adj-R^2	0.48	0.57	0.39	0.39	0.52
Obs	95	94	95	89	96

significant at: *10%, **5%, ***1%.

Using data from St. Louis FRED on per capita GDP by PPP to replicate AJR02 results.

they relate negatively to the instrument. Cohesiveness is only significant in two of the six however. Still, we would be concerned by the F statistics for the first stage, it is above ten only for column (4), suggesting we are using a weak instrument. I also added an extra panel to this table in which I remove from the regressions log of population density in 1500 and urbanization in 1500. We still observe low first stage F statistics for the most part and at the same time the instrument is not enough to push principles and cohesiveness out of the model most of the times.

Although my contention is that Principles do not require an instrument, because we have an exogenous shock in the colonization process, the next natural point is to see what occurs with the central results from AJR02 when we include this new model onto theirs. If Principles determines long-term development then Institutions should not be the cause of high current per capita GDP. However, to the extent we are using imperfect proxies for Principles, and both an imperfect measure and instrument for Institutions, we may see that adding the Principles' model lowers the importance of Institutions without completely taking over.

However, table 3.9 shows the opposite, most times the coefficient on Institutions is now larger than it was in the original paper. Further, Principles and cohesiveness are not as robust in this new set of IV regressions as they were before. Nevertheless, the measure of Principles used has significant direct or indirect weight on long-term wellbeing in five of the six regressions. Further, Panel C shows that by dropping urbanization and population density measures, which were never significant anyways, cohesiveness is also significantly impacting long-term wellbeing in four of the six regressions.

Finally, I tested the models adding Albouy's critique to it. Even though I have my own criticisms of the instrument beyond those of Albouy, in particular their coverage of only about two thirds of the observations, the relevant years for the data, and the simplistic treatment it gives to the concept of protection of property rights, Table 3.10 shows the effect of replicating AJR02 Table VIII only using the countries with original settler mortality data

Table 3.9: Replication Table VIII AJR02, with Model 1 (binary)

GDP per capita, Institutions, and Principles

Dependent variable is log GDP per capita by PPP in 1995						
	Expropriation Risk		Constraint on Exec. on 1990		Constraint on Exec. on Indep.	
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Second-Stage regressions					
Institutions	0.52***	1.36**	0.59*	0.53***	0.31**	0.66**
	(0.17)	(0.52)	(0.35)	(0.15)	(0.13)	(0.28)
Urbanization in 1500	0.006		0.06		0.02	
	(0.03)		(0.05)		(0.03)	
Log pop. density in 1500		-0.23		-0.01		-0.14
		(0.19)		(0.12)		(0.16)
Principles (binary)	0.42*	-0.27	-0.27	-0.22	0.67*	0.50
	(0.22)	(0.62)	(0.71)	(0.49)	(0.36)	(0.59)
Cohesive	0.33	-1.35	1.32**	0.95*	0.56	-0.71
	(0.43)	(1.14)	(0.55)	(0.51)	(0.47)	(1.00)
	Panel B: First-Stage regressions					
log settler mortality	-1.02***	-0.35**	-0.89*	-0.80***	-2.02***	-0.34
	(0.30)	(0.16)	(0.51)	(0.22)	(0.59)	(0.28)
Urbanization in 1500	-0.03		-0.11		-0.15	
	(0.05)		(0.08)		(0.09)	
Log pop. density in 1500		-0.01		-0.32		-0.08
		(0.16)		(0.22)		(0.28)
Principles (binary)	0.31	0.54	1.65**	1.79***	-1.87**	-1.75**
	(0.36)	(0.42)	(0.64)	(0.61)	(0.73)	(0.78)
Cohesive	0.35	0.98	-1.51	-1.46	-0.76	2.03
	(0.67)	(0.72)	(1.15)	(1.02)	(1.34)	(1.30)
F	9.08	7.11	3.85	12.68	6.00	2.83
Original F	19.52	12.21	3.19	19.23	13.63	6.28
Obs.	39	66	38	69	39	69
	Panel C: regressions Institutions & Principles only					
Institutions	0.53***	1.27**	0.69	0.53***	0.32**	0.62**
	(0.17)	(0.49)	(0.48)	(0.16)	(0.14)	(0.27)
Principles (binary)	0.40*	0.02	-0.56^{+}	-0.21^{+}	0.61*	0.67
	(0.21)	(0.54)	(0.98)	(0.51)	(0.35)	(0.53)
Cohesive	0.28	-0.55^{+}	0.91*	0.98**	0.41	0.24^{+}
	(0.39)	(0.90)	(0.54)	(0.39)	(0.45)	(0.77)
First stage F	19.45	9.58	4.40	21.22	8.20	8.42

$^{+}$ It has a significant positive effect through the instrument
significant at: *10%, **5%, ***1%.

plus six re-estimations done by Albouy (2012). In this case first stage F statistics go further down, making for a very weak instrument. More importantly, Institutions no longer help predict long-term development with this subset of the sample in four of the six regressions. At the same time, principles remain relevant for three of the six regressions (first or second stage) and cohesiveness in two.

However, the bigger concern here comes from the number of observations, at most 33 out of 102 former colonies. For former colonies other than 'neo-Europe', mortality rates exist for 45% of African countries, 28% of Asian countries, and 21% of American countries. This points to a non-random reason for the lack of data. This situation is further worsen when urbanization is included in the regression, causing the complete lost of the African data. The fact that the Principles model manages to stay relevant even in one of these regressions suggest it is quite robust.

Table 3.10: Replication Table VIII AJR02, with Model 1 (binary) and Albouy correction

GDP per capita, Institutions, and Principles

	Dependent variable is log GDP per capita by PPP in 1995					
	Expropriation Risk		Constraint on Exec. on 1990		Constraint on Exec. on Indep.	
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Second-Stage regressions					
Institutions	0.23	1.00*	0.30	0.51**	0.29	3.28
	(0.45)	(0.49)	(0.21)	(0.23)	(0.25)	(12.5)
Urbanization in 1500	0.03		0.12		0.10	
	(0.06)		(0.10)		(0.13)	
Log pop. density in 1500		-0.24		0.02		-0.34
		(0.16)		(0.18)		(1.37)
Principles (origin)	0.50	-0.81	0.41	0.40	0.66	-4.54
	(0.84)	(1.03)	(0.73)	(0.79)	(0.98)	(20.4)
Cohesive	1.35*	-0.91	2.05*	0.95	1.25	-8.64
	(0.80)	(1.18)	(1.02)	(0.88)	(1.08)	(37.1)
	Panel B: First-Stage regressions					
log settler mortality	-0.38	-0.11	-1.17**	-0.23	-1.09	0.57
	(0.35)	(0.19)	(0.54)	(0.30)	(0.82)	(0.39)
Urbanization in 1500	0.004		-0.35**		-0.32*	
	(0.09)		(0.10)		(0.15)	
Log pop. density in 1500		0.06		-0.27		0.26
		(0.19)		(0.29)		(0.38)
Principles (origin)	1.20	1.67**	1.71*	1.74	-0.13	0.69
	(0.78)	(0.77)	(0.86)	(1.23)	(1.31)	(1.59)
Cohesive	0.072	0.94	-4.72***	-2.24	-2.46	2.79
	(0.94)	(0.93)	(1.28)	(1.56)	(1.95)	(2.02)
F	5.85	7.47	9.41	4.93	6.21	2.42
Obs.	19	33	18	31	18	30
	Panel C: regressions Institutions & Principles only					
Institutions	-0.01	0.50	0.57	1.18	0.51	-0.32
	(0.73)	(0.99)	(0.81)	(2.33)	(0.99)	(0.61)
Principles (origin)	0.98	0.27^+	0.02	-1.16^+	0.45	1.34
	(1.33)	(1.87)	(1.90)	(5.30)	(1.54)	(1.32)
Cohesive	1.02	0.33	1.38	2.305	0.31	1.14
	(0.77)	(0.89)	(1.30)	(3.26)	(1.36)	(1.12)
First stage F	7.19	13.04	3.19	1.77	1.66	5.37

$^+$ It has a significant positive effect through the instrument

3.7 Conclusions

Although this type of cross country analyses have been largely given up lately, in favor of more closely followed local experiments, I do not think it can really be replaced by them. If we are looking to confirm a theory like that of institutions or this one for principles, we might be able to use local experiments, but without knowing what we are looking it will be hard or almost impossible to improve on the theory of institutions with local data only.

The principles' theory can better answer questions that have been raised over time about the theory of institutions. This is necessary to guide our research better both on local and global terms, to get answers that make sense to the questions that we raise. As I alluded in the previous paragraph, we will be hard pressed to find the right answers by looking for them under the wrong assumptions, I hope this new theory can bring us closer to the right ones and accelerate the advancement of our knowledge.

This could be a very rich topic of research for the coming years. There are many related issues that need to be explored to improve our understanding of how long term development occurs. This includes improvement on this paper itself, such as data collection and probably better minds providing more answer, but also tests to the assumptions made and the theories surrounding it. On doing this research I suffered greatly with my own limitations on a very vast body of knowledge.

Bibliography

Aaronson, Daniel, "Price pass-through and the minimum wage," *Review of Economics and statistics*, 2001, *83* (1), 158–169.

_ , Eric French, and James MacDonald, "The minimum wage, restaurant prices, and labor market structure," *Journal of Human Resources*, 2008, *43* (3), 688–720.

_ , _ , Isaac Sorkin, and Ted To, "Industry dynamics and the minimum wage: a putty-clay approach," *International Economic Review*, 2018, *59* (1), 51–84.

_ , Kyung-Hong Park, Daniel Sullivan et al., "Explaining the decline in teen labor force participation," *Chicago Fed Letter*, 2007, *235.*

Abowd, John M, Francis Kramarz, Thomas Lemieux, and David N Margolis, "Minimum wages and youth employment in France and the United States," Technical Report, National Bureau of Economic Research 1997.

Acemoglu, Daron, Simon Johnson, and James A Robinson, "The colonial origins of comparative development: An empirical investigation," *American economic review*, 2001, *91* (5), 1369–1401.

_ , _ , and _ , "Reversal of fortune: Geography and institutions in the making of the modern world income distribution," *The Quarterly journal of economics*, 2002, *117* (4), 1231–1294.

_ , _ , and _ , "The colonial origins of comparative development: An empirical investigation: Reply," *American Economic Review*, 2012, *102* (6), 3077–3110.

Albouy, David Y, "The colonial origins of comparative development: an empirical investigation: comment," *American Economic Review*, 2012, *102* (6), 3059–76.

Allegretto, Sylvia A, Arindrajit Dube, and Michael Reich, "Do minimum wages really reduce teen employment? Accounting for heterogeneity and selectivity in state panel data," *Industrial Relations: A Journal of Economy and Society*, 2011, *50* (2), 205–240.

Alston, Eric, Lee J Alston, Bernardo Mueller, and Tomas Nonnenmacher, *Institutional and organizational analysis: concepts and applications*, Cambridge University Press, 2018.

Alston, Lee J, Marcus André Melo, Bernardo Mueller, and Carlos Pereira, "Why Countries Transition? The Case of Brazil, 1964–2016," *Atlantic Economic Journal*, 2016, *44* (2), 197–224.

Alzúa, María Laura, Guillermo Cruces, and Laura Ripani, "Welfare Programs and Labor Supply in Developing Countries: experimental evidence from Latin America," *Journal of Population Economics*, 2013, *26* (4), 1255–1284.

Attanasio, Orazio, Erich Battistin, Emla Fitzsimons, and Marcos Vera-Hernandez, "How effective are Conditional Cash Transfers? Evidence from Colombia," 2005.

Banerjee, Abhijit V, Rema Hanna, Gabriel E Kreindler, and Benjamin A Olken, "Debunking the stereotype of the lazy Welfare Recipient: Evidence from Cash Transfer Programs," *The World Bank Research Observer*, 2017, *32* (2), 155–184.

Bank, World, "World Bank Open Data," 2018.

Barattieri, Alessandro, Susanto Basu, and Peter Gottschalk, "Some evidence on the importance of sticky wages," *American Economic Journal: Macroeconomics*, 2014, *6* (1), 70–101.

Basker, Emek and Muhammad Taimur Khan, "Does the minimum wage bite into fast-food prices?," *Journal of Labor Research*, 2016, *37* (2), 129–148.

Bayle, Constantino, *Los cabildos seculares en la América española*, Sapientia, 1952.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, "How much should we trust Differences-in-differences Estimates?," *The Quarterly Journal of Economics*, 2004, *119* (1), 249–275.

Besley, Timothy and Torsten Persson, "The origins of state capacity: Property rights, taxation, and politics," *American Economic Review*, 2009, *99* (4), 1218–44.

_ **and** _ , *Pillars of prosperity: The political economics of development clusters*, Princeton University Press, 2011.

Brauw, Alan De, Daniel O Gilligan, John Hoddinott, and Shalini Roy, "Bolsa Família and Household Labor Supply," *Economic Development and Cultural Change*, 2015, *63* (3), 423–457.

Campa, Arthur L, "The churchmen and the Indian languages of New Spain," *The Hispanic American Historical Review*, 1931, *11* (4), 542–550.

Card, David, "Do minimum wages reduce employment? A case study of California, 1987–89," *ILR Review*, 1992, *46* (1), 38–54.

_ **and Alan B Krueger**, "Minimum wages and employment: A case study of the fast food industry in New Jersey and Pennsylvania," Technical Report, National Bureau of Economic Research 1993.

Clemens, Jeffrey and Michael Wither, "The minimum wage and the Great Recession: Evidence of effects on the employment and income trajectories of low-skilled workers," *Journal of Public Economics*, 2019.

Cogan, John, "Fixed costs and labor supply," 1980.

Consejo de las Indias, *Recopilación de Leyes de los Reynos de las Indias - Tomo Primero*, Vol. 4, Consejo de la Hispanidad, 1943.

Currie, Janet and Bruce Fallick, "The minimum wage and the employment of youth: Evidence from the NLSY," Technical Report, National Bureau of Economic Research 1993.

de Barradas, José Pérez, *Los mestizos de América* number 1610, Espasa-Calpe, 1976.

de Las Casas, Bartolomé, "Historia de las Indias, lib. III, caps. 3 and 4."

Dell, Melissa, "The persistent effects of Peru's mining mita," *Econometrica*, 2010, *78* (6), 1863–1903.

_ , Nathan Lane, and Pablo Querubin, "The historical state, local collective action, and economic development in Vietnam," *Econometrica*, 2018, *86* (6), 2083–2121.

Deus, Sublimis, "Papst Paul III., Bulle Sublimis Deus. 9. Juni 1537," *Michael Sievernich ua (Hg.), Conquista und Evangelisation*, *500*, 475–476.

Dirección de Presupuesto de Chile, "Gasto Gobierno Central como porcentaje del PIB," 2019.

Dube, Arindrajit, T. William Lester, and Michael Reich, "Minimum wage effects across State borders: Estimates using contiguous counties," *The Review of Economics and Statistics*, 2010, *92* (4), 945–964.

_ , T William Lester, and Michael Reich, "Minimum wage shocks, employment flows, and labor market frictions," *Journal of Labor Economics*, 2016, *34* (3), 663–704.

Elliott, John Huxtable, *Empires of the Atlantic world: Britain and Spain in America, 1492-1830*, Yale University Press, 2007.

Engerman, Stanley L and Kenneth L Sokoloff, "Factor endowments, inequality, and paths of development among new world economics," Technical Report, National Bureau of Economic Research 2002.

Errington, Joseph, *Linguistics in a colonial world: A story of language, meaning, and power*, John Wiley & Sons, 2008.

Estrugo, José M, *Los sefardíes*, Vol. 2, Editorial Renacimiento, 2002.

Federal Rerseve Bank of St. Louis, "Purchasing Power Parity Converted GDP per Capita," 2019.

Fiszbein, Ariel and Norbert R Schady, *Conditional Cash Transfers: reducing present and future Poverty*, The World Bank, 2009.

Garganta, Santiago and Leonardo Gasparini, "The impact of a Social Program on Labor informality: The case of AUH in Argentina," *Journal of Development Economics*, 2015, *115*, 99–110.

Gertler, Paul, "Do Conditional Cash Transfers improve Child Health? Evidence from PROGRESA's Control Randomized Experiment," *American Economic Review*, 2004, *94* (2), 336–341.

Giuliano, Laura, "Minimum wage effects on employment, substitution, and the teenage labor supply: Evidence from personnel data," *Journal of Labor Economics*, 2013, *31* (1), 155–194.

Hanke, Lewis, *The Spanish Struggle for Justice in the Conquest of America*, University of Pennsylvania Press, 1949.

Herrera, Rodrigo, Osvaldo Larra naga, and Amanda Telias, "La Ficha de Protección Social," 2010.

Insua, Pedro, *1492. España contra sus fantasmas: Prólogo de María Elvira Roca Barea*, Editorial Ariel, 2018.

Irarrázabal, Ignacio, Alejandra Candia, Rodrigo Castro, Germán Codina, Rodrigo Delgado, Rodrigo Herrera, Julio Guzmán, Osvaldo Larra naga, Cecilia Ormazábal, Mahia Saracostti, Mónica Titze, and Salvador Valdés, "Comite de Expertos Ficha de Protección Social," 2010.

Kagan, Richard L, Fernando Marías, and Fernando Marías Franco, *Urban images of the Hispanic world, 1493-1793*, Yale University Press, 2000.

Lam, David and Suzanne Duryea, "Effects of schooling on fertility, labor supply, and investments in children, with evidence from Brazil," *Journal of Human Resources*, 1999, pp. 160–192.

Licona, Gonzalo Hernández, Thania De la Garza, Janet Zamudio, and Iliana (coords.) Yaschine, "El Progresa-Oportunidades-Prospera, a 20 aos de su creación," 2019.

Lummis, Charles F, Arturo Cuyás Armengol, and Rafael Altamira, *Los exploradores españoles del siglo XVI*, Ediciones Palabra, 1989.

Maddison, Angus, "The Maddison-Project," *línea] http://www. ggdc. net/maddison/maddison–project/home. htm*, 2013, *1*, 14.

Maluccio, John and Rafael Flores, *Impact evaluation of a Conditional Cash Transfer Program: The Nicaraguan Red de Protección Social*, Intl Food Policy Res Inst, 2005.

Meer, Jonathan and Jeremy West, "Effects of the minimum wage on employment dynamics," *Journal of Human Resources*, 2016, *51* (2), 500–522.

Ministerio de Desarrollo Social, "Encuesta CASEN 1998," 1999.

_ , "Encuesta CASEN 2000," 2001.

_ , "Encuesta CASEN 2003," 2004.

_ , "Encuesta CASEN 2006," 2007.

_ , "Encuesta CASEN 2009," 2010.

_ , "Encuesta CASEN 2011," 2012.

_ , "Encuesta CASEN 2013," 2014.

_ , "Encuesta CASEN 2015," 2016.

Neumark, David and William L Wascher, *Minimum wages*, MIT press, 2008.

_ **and William Wascher**, "The effects of minimum wages on teenage employment and enrollment: Evidence from matched CPS surveys," Technical Report, National Bureau of Economic Research 1995.

_ **, Mark Schweitzer, and William Wascher**, "Minimum wage effects throughout the wage distribution," *Journal of Human Resources*, 2004, *39* (2), 425–450.

Newman, John Henry, George P Landow, Juan Enrique Newman, Frank Miller Turner, Martha McMackin Garland, and Sara Castro-Klaren, *The idea of a university*, Yale University Press, 1996.

Pedace, Roberto and Stephanie Rohn, "The impact of minimum wages on unemployment duration: Estimating the effects using the Displaced Worker Survey," *Industrial Relations: A Journal of Economy and Society*, 2011, *50* (1), 57–75.

Pérez, Demetrio Ramos, *Historia de la colonización española en América*, Pegaso, 1947.

Powell, David, "Synthetic Control Estimation Beyond Case Studies: Does the Minimum Wage Reduce Employment?," 2017.

Ribas, Rafael and Fábio Veras Soares, "Is the effect of Conditional Transfers on Labor Supply negligible everywhere?," 2011.

Schultz, T Paul, "Returns to women's education," *Women's education in developing countries: Barriers, benefits, and policies*, 1993, pp. 51–99.

_ , "School Subsidies for the Poor: Evaluating the Mexican PROGRESA Poverty Program," *Journal of Development Economics*, 2004, *74* (1), 199–250.

Skoufias, Emmanuel and Vincenzo Di Maro, "Conditional Cash Transfers, Adult Work incentives, and Poverty," *The Journal of Development Studies*, 2008, *44* (7), 935–960.

Sorkin, Isaac, "Are there long-run effects of the minimum wage?," *Review of economic dynamics*, 2015, *18* (2), 306–333.

Strauss, John and Duncan Thomas, "Human resources: Empirical modeling of household and family decisions," *Handbook of development economics*, 1995, *3*, 1883–2023.

Subsecretaría de Seguridad Social, "Estadísticas Anuales, 2016," 2018.

The World Bank, "The State of Social Safety Nets 2018," *Washington, DC: World Bank. doi:10.1596/978-1-4648-1254-5. License: Creative Commons Attribution CC BY 3.0 IGO*, 2018.

Thompson, Jeffrey P, "Using local labor market data to re-examine the employment effects of the minimum wage," *ILR Review*, 2009, *62* (3), 343–366.

Vaghul, Kavya and Ben Zipperer, "Historical state and sub-state minimum wage data," *Washington Center for Equitable Growth*, 2016.

Zavala, Silvio A, *La Encomienda Indiana*, Madrid, [imprenta helénica, 1935.

Zavodny, Madeline, "The effect of the minimum wage on employment and hours," *Labour Economics*, 2000, *7* (6), 729–750.

Appendix A

Appendix to Chapter I

Having children can be so transformative that we may doubt a comparison between women with and without children. Correspondingly, I tested alternative definitions of treatment and control groups that bridge this gap. However, these definitions have limitations by design, wich lead me to create two alternative definitions, so that they can confirm and complete each other. Non mothers are completely ignored in what follows, with control being conformed instead by mothers that we can logically expect to be less responsive to the system's reform, as they can expect lesser gains.

One of these alternative definitions relies on number of children to distinguish treatment from control, and the other uses the age of the youngest child for the same purpose.

Number of Children

The subsidy under study is given in per capita basis, one per child (and one for the mother). This means that each extra child increases the transfer significantly. The second child increases the monthly transfer by one third, third child increases its overall value by twenty five percent, and so on. The chosen categorization in column three of table A.1 means that the average number of children of the control group is 1.47, and for treatment is 3.34. This implies that the value of the monthly transfer for the treatment group is 76% higher, assuming all these children are complying with the conditions for the transfer. We would expect therefore that, when the subsidy is reformed in 2007, and turned into an entitlement, women with more children will be more likely to apply to the program, since it benefits them more.

Table A.1: Mothers by number of children

Children	Frequency	Status
1	138,983	control
2	124,275	control
3	54,543	treatment
4	14,887	treatment
5	3,691	treatment
6	1,038	treatment
7	330	treatment
8	143	treatment
9	32	treatment
+10	20	treatment

This allows for some important differences between the original treatment and control group to be abridged. First, and most important, all are mothers. But also other differences are reduced: years of schooling, albeit marginally; percentage receiving the subsidy; and labor supply related variables. In fact, only the percentage of mothers that have no spouse is considerably different between treatment and control still. Nevertheless, figure A.1 puts some doubt on whether the parallel trends assumption holds for this model.

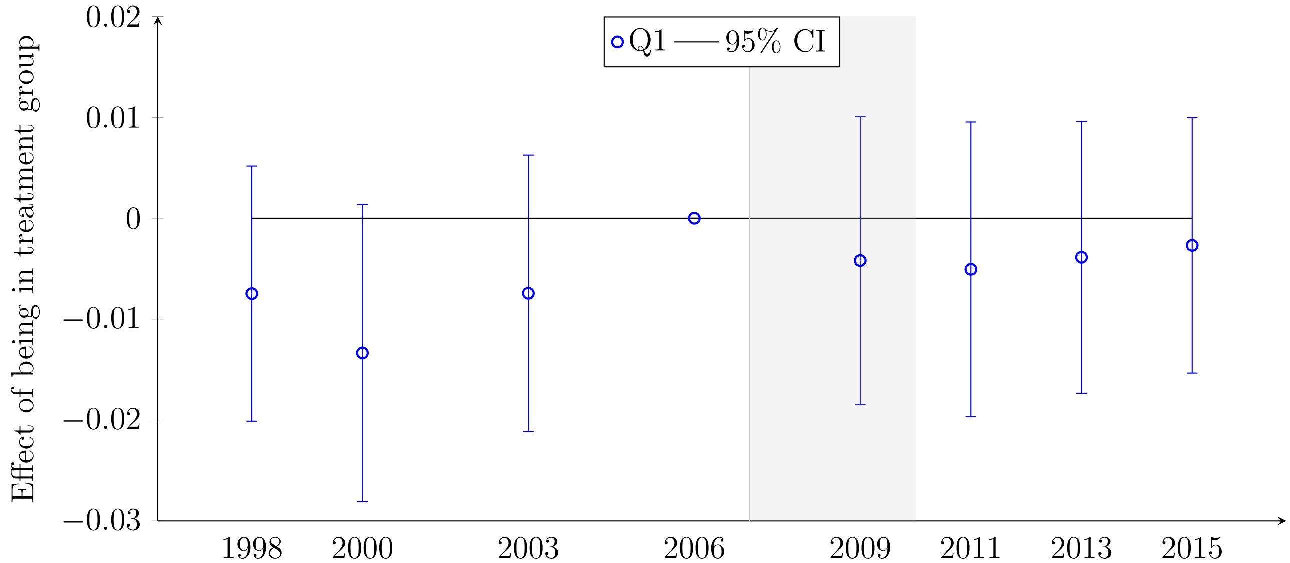
There is another penalty to pay for this choice. Almost no women under twenty five years old has three or more children, making the analysis of that important demographic impossible using this definition for treatment and control.

Table A.2: Descriptive Statistics for treatment/control

Year	Obs	Age	Years Educ.	# children	First Mom	% Single	% SUF	LF partic.	Work Income*	Work Hours
Treatment										
1998	31,969	37.3	9.6	3.42	21.9	12.7	17.7	40.2	739,836	14.8
2006	43,159	38.6	10.6	3.38	22.3	16.0	15.0	49.9	552,467	18.2
2015	40,025	37.8	11.4	3.26	22.0	26.3	28.2	57.8	411,245	20.5
Control										
1998	3,495	33.6	10.6	1.50	23.9	24.4	9.9	48.0	607,065	18.3
2006	5,050	35.0	11.1	1.48	24.3	29.8	10.3	56.3	497,261	20.9
2015	7,109	35.3	12.1	1.44	24.7	38.9	19.5	64.8	405,913	24.0

*refers to values for mothers working at least 1 weekly hour, in real value

Figure A.1: Test for Parallel Trends - By Number of children



Youngest Child

To address the important limitations of the previous approach I explored an additional dimension along which the benefits of the policy differ, age of the youngest child and essentially the same economical argument. Mothers with younger children can expect to receive the subsidy for a longer time, so they benefit more from applying to it after the reform makes it more accessible.

Indeed, when mothers whose youngest child is at most eleven years old, the average age of the youngest child for mothers in the treatment group is 4.7, compared to 16.1 for control. That means that, even at a ten percent discount rate, the program is almost three times more beneficial for the mother of the younger child, if considered till the age of eighteen, and still fifty percent more valuable when estimated till twenty four years of age. Furthermore, all things equal, a mother with a younger child is more likely to have more children in the program as well, adding to the difference.

As table A.3 shows, these groups are more comparable in terms of numbers of children, single status, and income; and are also not too unbalanced in labor force participation, education, and hours worked. As we would expect, mothers with younger children are significantly more likely to participate in the program. Additionally, the parallel trends assumption is better supported by this approach, as is shown in figure A.2.

The limitation of this third definition is that again all younger mothers are fit squarely in one category, although in this case is treatment. This means that we will be using as control for them older mothers exclusively.

Table A.3: Descriptive Statistics for treatment/control

Year	Obs	Age	Years Educ.	# children	First Mom	% Single	% SUF	LF partic.	Work Income*	Work Hours
Treatment										
1998	31,969	32.7	10.4	2.13	23.1	20.6	13.8	43.6	602,841	16.0
2006	43,159	33.4	11.2	2.02	23.3	26.1	13.1	52.3	503,820	18.9
2015	40,025	33.1	12.2	1.86	23.7	35.7	23.8	60.9	407,706	21.7
Control										
1998	3,495	43.2	9.5	1.92	23.9	22.3	5.3	52.5	770,764	21.7
2006	5,050	43.4	10.3	1.81	24.4	27.3	6.5	60.4	519,115	23.4
2015	7,109	43.8	11.3	1.63	24.9	38.0	13.0	69.9	401,836	27.5

*refers to values for mothers working at least 1 weekly hour, in real value

Figure A.2: Test for Parallel Trends - By Number of children

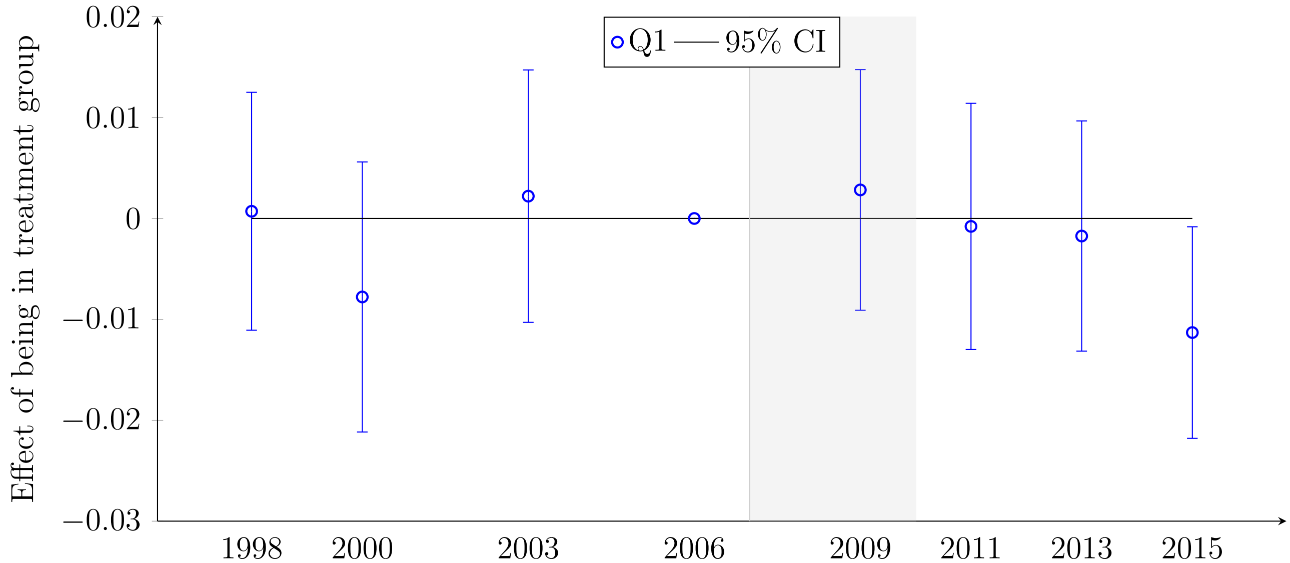


Table A.4: Labor Supply Response Overall - Number of children

	Unconditional			**Conditional**		
	Labor Force partic.	Hours worked	Hours worked others	Hours worked (logs)	Prob. overtime	Hourly wage
	(1)	(2)	(3)	(4)	(5)	(6)
treated	0.005	0.114	1.008**	-0.008	0.008	-0.025
[s.e.]	[0.004]	[0.238]	[0.373]	[0.016]	[0.009]	[0.023]
mean	*49.9%*	*18.2*	*38.8*	*40.0*	*13.3%*	*$2,704*
Observations	**263,006**		**188,260**		**103,917**	

significant at: ***0.1%, **1%, *5%, +10%

All regressions include demographic controls. Unconditional outcomes control for whether the subject has work experience while conditional outcomes control for time in current job. The means used correspond to the year 2006.

Results

Tables A.4 through A.11 show for both these approaches the same analysis done in the paper for the chosen treatment/control definition. To note particularly, these results confirm the negative response by young mothers to treatment and the positive response by more educated mothers. Furthermore, although the result is not shown here the approach by youngest child also confirms that older mothers with spouse marginally increase their labor force participation, even though the approach by number of children suggests zero response.

Table A.5: Labor Supply Response Overall - Youngest child

	Unconditional			**Conditional**		
	Labor Force partic.	Hours worked	Hours worked others	Hours worked (logs)	Prob. overtime	Hourly wage
	(1)	(2)	(3)	(4)	(5)	(6)
treated	0.003	-0.343	-1.056**	-0.032**	0.012	0.038
[s.e.]	[0.004]	[0.216]	[0.333]	[0.012]	[0.009]	[0.027]
mean	*52.3%*	*18.9*	*34.0*	*40.2*	*12.1%*	*$2,566*
Observations	**263,006**		**188,260**		**103,917**	

significant at: ***0.1%, **1%, *5%, $^{+}$10%

All regressions include demographic controls. Unconditional outcomes control for whether the subject has work experience while conditional outcomes control for time in current job. The means used correspond to the year 2006.

Table A.6: Labor Supply Response by Age group - Number of children

	Unconditional			**Conditional**		
	Labor Force partic.	Hours worked	Hours worked others	Hours worked (logs)	Prob. overtime	Hourly wage
	(1)	(2)	(3)	(4)	(5)	(6)
treated						
18 to 24	-0.035^{+}	-0.536	-1.933***	-0.085	-0.176***	0.206
[s.e.]	[0.018]	[0.569]	[0.231]	[0.087]	[0.044]	[0.15]
mean	*33.9%*	*10.6*	*38.3*	*38.9*	*29.0%*	*$2,589*
25 to 50	0.003	0.086	1.033**	-0.014	0.006	-0.022
[s.e.]	[0.004]	[0.246]	[0.381]	[0.017]	[0.009]	[0.024]
mean	*50.1%*	*18.3*	*38.8*	*40.1*	*13.1%*	*$2,705*
Observations	**263,006**		**188,260**		**103,917**	

significant at: ***0.1%, **1%, *5%, $^{+}$10%

All regressions include demographic controls. Unconditional outcomes control for whether the subject has work experience while conditional outcomes control for time in current job. The means used correspond to the year 2006.

Table A.7: Labor Supply Response by Age group - Youngest child

	Unconditional			**Conditional**		
	Labor Force partic.	Hours worked	Hours worked others	Hours worked (logs)	Prob. overtime	Hourly wage
	(1)	(2)	(3)	(4)	(5)	(6)
treated						
18 to 24	-0.053	-4.359***	5.113	-0.251	-0.179	-0.527^{+}
[s.e.]	[0.065]	[0.045]	[929.543]	[0.282]	[0.168]	[0.271]
mean	*41.4%*	*13.3*	*23.3*	*40.1*	*11.6%*	*$1,385*
25 to 50	0.006^{+}	-0.296	-1.039**	-0.024^{+}	0.016^{+}	0.03
[s.e.]	[0.004]	[0.225]	[0.339]	[0.013]	[0.010]	[0.027]
mean	*54.3%*	*19.9*	*35.8*	*40.3*	*12.1%*	*$2,703*
Observations	**263,006**		**188,260**		**103,917**	

significant at: ***0.1%, **1%, *5%, $^{+}$10%

All regressions include demographic controls. Unconditional outcomes control for whether the subject has work experience while conditional outcomes control for time in current job. The means used correspond to the year 2006.

Table A.8: Labor Supply Response by Single/Couple - Number of children

	Unconditional			**Conditional**		
	Labor Force partic.	Hours worked	Hours worked others	Hours worked (logs)	Prob. overtime	Hourly wage
	(1)	(2)	(3)	(4)	(5)	(6)
treated						
single	-0.0002	-0.484	-	-0.014	0.012	0.004
[s.e.]	[0.01]	[0.612]	-	[0.027]	[0.017]	[0.033]
mean	*76.9%*	*28*	-	*40*	*13%*	*$2,862*
in couple	0	0.139	1.008**	-0.019^{+}	0.001	-0.029*
[s.e.]	[0.003]	[0.178]	[0.373]	[0.011]	[0.006]	[0.014]
mean	*44.8%*	*16.3*	*46.2*	*40.1*	*13.9%*	*$2,230*
Observations	**263,006**		**188,260**		**103,917**	

significant at: ***0.1%, **1%, *5%, $^{+}$10%

All regressions include demographic controls. Unconditional outcomes control for whether the subject has work experience while conditional outcomes control for time in current job. The means used correspond to the year 2006.

Table A.9: Labor Supply Response by Single/Couple - Youngest child

	Unconditional			Conditional		
	Labor Force partic.	Hours worked	Hours worked others	Hours worked (logs)	Prob. overtime	Hourly wage
	(1)	(2)	(3)	(4)	(5)	(6)
treated						
single	-0.011^{+}	-0.537	-	-0.050**	0.014	0.066**
[s.e.]	[0.006]	[0.431]	-	[0.018]	[0.015]	[0.024]
mean	*70%*	*25.5*	-	*39.7*	*11.2%*	*$2,925*
in couple	0.006*	-0.085	-1.056**	-0.020*	0.011*	0.027*
[s.e.]	[0.003]	[0.133]	[0.333]	[0.009]	[0.005]	[0.013]
mean	*46.1%*	*16.5*	*46*	*41.4*	*13.6%*	*$1,886*
Observations	**263,006**		**188,260**		**103,917**	

significant at: ***0.1%, **1%, *5%, $^{+}$10%

All regressions include demographic controls. Unconditional outcomes control for whether the subject has work experience while conditional outcomes control for time in current job. The means used correspond to the year 2006.

Table A.10: Labor Supply Response by Education Attainment - Number of children

	Unconditional			Conditional		
	Labor Force partic.	Hours worked	Hours worked others	Hours worked (logs)	Prob. overtime	Hourly wage
	(1)	(2)	(3)	(4)	(5)	(6)
treated						
$\leq$ *High school*	0.002	0.135	0.861^{+}	-0.0001	0.011	-0.037
[s.e.]	[0.005]	[0.264]	[0.450]	[0.020]	[0.011]	[0.027]
mean	*44.7%*	*16*	*38.3*	*40.4*	*9.8%*	*$4,573*
$>$ *High school*	0.014*	0.093	-2.767***	-0.022	0.005	0.007
[s.e.]	[0.006]	[0.312]	[0.104]	[0.016]	[0.009]	[0.020]
mean	*69.3%*	*26.5*	*40.6*	*39.9*	*14.8%*	*$1,865*
Observations	**263,006**		**188,260**		**103,917**	

significant at: ***0.1%, **1%, *5%, $^{+}$10%

All regressions include demographic controls. Unconditional outcomes control for whether the subject has work experience while conditional outcomes control for time in current job. The means used correspond to the year 2006.

Table A.11: Labor Supply Response by Education Attainment - Youngest child

	Unconditional			**Conditional**		
	Labor Force partic.	Hours worked	Hours worked others	Hours worked (logs)	Prob. overtime	Hourly wage
	(1)	5 (2)	(3)	(4)	(5)	(6)
treated						
$\leq$ *High school*	0.004	-0.197	-0.948*	-0.034*	0.008	0.043
[s.e.]	[0.004]	[0.239]	[0.383]	[0.015]	[0.011]	[0.033]
mean	*46.7%*	*16.7*	*34*	*40.4*	*14%*	*$1,616*
$>$ *High school*	-0.004	-0.819***	6.025	-0.044**	0.009	0.040*
[s.e.]	[0.005]	[0.114]	[2522]	[0.013]	[0.008]	[0.018]
mean	*70%*	*25.8*	*33.8*	*39.9*	*8.1%*	*$4,437*
Observations	**272,609**		**195,743**		**107,693**	

significant at: ***0.1%, **1%, *5%, $^{+}$10%

All regressions include demographic controls. Unconditional outcomes control for whether the subject has work experience while conditional outcomes control for time in current job. The means used correspond to the year 2006.