Border enforcement, aid and migration

Manuela Angelucci

A Dissertation submitted to the Department of Economics
in partial fulfilment of the requirements for the degree of

Doctor of Philosophy
University College London

London

2004

Abstract

This thesis addresses the issue of policy effects on domestic and international migration, considering in particular the case of Mexican migration. The first essay investigates the effect of U.S. border enforcement on the net flow of Mexican undocumented migration. Such effect is theoretically ambiguous, given that increases in border controls deter prospective migrants from crossing the border illegally, but lengthen the U.S. permanence of current ones. It estimates border enforcement's net impact on migration inflow using a sample of potential and current illegal migrants. U.S. border enforcement significantly reduces the net flow of undocumented migration. However, the reduction in net flow is more than half the size of the decrease in inflow.

The second essay models the short and medium-run impact of aid on migration, considering alternatively the effect of unconditional and conditional cash transfers to financially constrained households. Data from the evaluation of a Mexican development program, Progresa, are used to estimate the effect of the grant on migration. The empirical analysis shows that the program is associated with an increase in international migration, which is also a positive function of the potential transfer size. Conditional grants in the form of secondary school subsidies reduce the short-term migration probability. Progresa does not seem to increase medium-term migration.

The final chapter reviews the approaches employed to estimate Treatment on the Treated Effects (TTEs) using experimental data in the presence of non-compliers. It discusses the types of parameters that can be identified using the Progresa data. It uncovers new parameters that have not been estimated so far, based on the fact that a group of eligible households did not receive the program transfer in the initial stages of its implementation. It proposes alternative estimating procedures to identify counterfactuals in the presence of non-compliers for users of the Progresa data. It complements the theoretical part with an empirical application by estimating the effect of Progresa on school enrolment.

Contents

List of Tables

List of Figures

Acknowledgements

It is a truth universally acknowledged that this thesis would have not been written without the help, guidance and support of the following people and institutions, to which I am most grateful.

First of all I would like to thank my supervisors, Orazio Attanasio and Costas Meghir. Among other things, they tried to provide me with a set of theoretical and empirical tools to identify, develop and test interesting research ideas and to teach me how to analyse them critically and concisely.

Many UCL faculty members, IFS staff members and PhD students provided excellent comments to my papers, in particular Sami Berlinski, Richard Blundell, Pedro Carneiro, Andrew Chesher, Marco Cozzi, Giacomo De Giorgi, Emla Fitzsimons, Mario Fiorini, Antonio Guarino, Nicola Pavoni and Cristina Santos, whom I thank also for the "technical assistance", but especially for her support.

Jeff Smith has been my "honorary" supervisor. Not only did I learn much from his lectures and presentations and from our conversations; he also showed an enormous degree of support during the job market and throughout my (few) career choices. Most of all, he has been the living proof that also economists can be fun and decent human beings. I am most indebted to him.

Gian Luigi Albano is single-handedly responsible for my material and psychological survival during the last 12 months. I thank him for the many nights he fed me when coming home late, for shedding light on the many shades and the many sides of several economic issues, but primarily for our illuminating dialogues on life, the universe and everything. I also thank him for his patience, his care and for reading innumerable drafts of my papers.

I took my first steps in the realm of economics while working with Saul Estrin and Mario Nuti at London Business School. I saw for the first time how a research idea is conceived and a paper is developed; I did my very first presentation at the Bureau for Economic Analysis in Moscow.

I benefited immensely from the many conversations on economics and human behaviour, poetry and music with Adam Szeidl, with whom I also shared the excitement and difficulties of the job market, as well as a couple of "Hu Bazmeg".

Several institutions supported me financially: the Italian National Research Council (CNR), the Centre for Economics Studies "Luigi Einaudi", the Royal Economic Society and the University of Rome La Sapienza, which provided me with a research grant to complete the third chapter of this thesis under the supervision of Marina Capparucci.

I would also like to thank the people I met in these last few years, and of course the "old guard"; the friends who helped me stay sane and reminded me that there is a life beyond the PhD. Listing names is unnecessary: they know who they are and how much they mean to me.

Finally, I would like to thank my family: those who supported me silently and those who talked too much. The ones who never visited and those who would never leave. My younger cousins, who, after meeting their first economist, whispered "Now we know why she hasn't got married yet!". They may not understand what I am doing, but they are always there for me.

Declaration

1. No part of this thesis has been presented to any University for any degree.

Manuela Angelucci

Chapter 1

Introduction

The first two chapters of this thesis study the impact that different policies have on both international and domestic migration. In particular, the questions I try to answer are whether enforcing the border of the destination country reduces the net flow of illegal immigration from the country of origin; whether there are alternative means to achieve this objective, and, in particular, how aid programs affects current and future domestic and international migration. The final part of my thesis deals with the estimation of treatment on the treated effects (TTEs) in the experimental data set used in the second chapter. I present a new set of identifying assumptions, different from the ones normally made by the users of these data, and I uncover several new parameters that have clear policy relevance.

The focus of the empirical analysis is on Mexican migration because of its interesting features. Mexico is a country where both domestic and international labour migration occurs. Given the geographic proximity and the large wage differential, almost all international migrants are directed to the United States. The bulk of Mexican migration to the U.S. is illegal and composed of highly mobile temporary movers, who may migrate several times over the life cycle. The size of the phenomenon is considerable: according to the U.S. Immigration and Naturalization Service, some 5 million undocumented Mexicans lived in the United States in 2000, 69 percent of the overall illegal population. Nevertheless, the theoretical conclusions can be easily extended to different sets of countries.

1.1 U.S. border enforcement and the net flow of Mexican illegal migration

Especially since the mid 1980s, the U.S. government has been exerting considerable effort to discourage illegal immigration. The 1986 Immigration Reform and Control Act has introduced employer's sanctions for hiring undocumented workers and devoted more resources to the apprehension of aliens. The tougher enforcement has resulted in substantial increases in border patrol. Aliens apprehended trying to cross the southern U.S. border are in most cases not prosecuted, given the high associated administrative costs. They are instead driven back to the Mexican side of the border, from which they normally soon attempt re-entry.

To date, there is a small but growing literature that tries to evaluate the effect on border enforcement on the inflow of illegal migration. However, these studies suffer from two limitations: first, they try to infer the deterrent effect of border controls on prospective migrants by looking at changes in border apprehensions. The relationship between the latter and the size of migration is not straightforward, because of the possibility of repeated border-crossing attempts within a single trip. Moreover, in order to assess the policy impact on the size of the undocumented U.S. residents, one has to understand how the behaviour of current illegal migrants is affected by changes in border enforcement. Tougher policing of the border translates in higher migration costs. Higher costs may reduce the number of people entering the United States illegally, but at the same time it is likely to lengthen migration spells, because it takes longer to reap a positive return from the trip. Also, if undocumented migrants who are currently in the U.S. anticipate future tougher enforcement, they may increase the duration of the current migration knowing that future trips will be more costly. Since higher border policing reduces both the inflow and the outflow of illegal migration, its effect on migration net inflow, hence stock, is ambiguous and it depends on the size of potential and current undocumented migrant populations and on how sensitive to changes of border patrol each group's behaviour is.

The empirical analysis estimates the effect that changes in the intensity of border enforcement between 1972 and 1993 have on both the U.S. inflow and outflow of illegal Mexican migrants modelling individual transitions between two states: potential and current illegal migrant. The estimations are undertaken using a unique data set that merges aggregate border enforcement linewatch data with information on a sample of illegal potential and de facto migrants from the Mexican Migration Project. The shortcomings associated with the particular design of the survey are extensively discussed. The econometric model accounts for the endogeneity of border controls using Drugs Enforcement Administration budget as instrumental variable. After computing the size of the illegal undocumented inflow, the estimated marginal effects are used to measure the effect of a marginal change in enforcement on the number of individuals who are respectively deterred from migrating illegally and from returning to Mexico from an undocumented U.S. migration. I show that the deterrent effect is small but significant, in line with the existing literature. When the outflow effect is included as well, the size of the net impact is still positive, but halved. Hence, estimates of the effectiveness of border patrol based on its reduction in inflow are grossly overestimating its true effect. I provide back-of-the envelope measures of the marginal cost of reducing the stock of migration by one person, which amounts to a few hundred dollars. Furthermore, an additional cost of this policy is its contribution to the creation of a more permanent undocumented resident population.

1.2 Aid and migration: an analysis of the impact of Progresa on the timing and size of Mexican migration

In the current analysis I study the impact of aid policies on current and future domestic and international migration. In particular, I consider the effects of an unconditional transfer and of a conditional schooling subsidy. Transfers to prospective migrants reduce the economic disparities between home and the destination location, lowering the incentives to leave. However, if individuals or households are credit constrained, the grants may be used to finance new migrations. The latter fact may have sizeable effects, given that the imperfection of capital markets is one acute problem faced by indigent families in developing and middle-income countries. The conditional transfer that requires recipients to stay in the home country may reduce migration in the short-term. Nevertheless, individuals may revert to migrating once the subsidy ends. Future migration levels may even be higher than in the zero-aid scenario, if recipients save in order to fund future trips.

A development policy that has all the aforementioned features is Progresa, a program that targets poor Mexican rural households. Among the various components of the program, there are a (smaller) unconditional nutrition support grant, and then some (cumulatively larger) transfers conditional upon attendance of the last four grades of primary school and first three grades of secondary school, all paid bi-monthly. The subsidy to primary school attendance is roughly equivalent to an unconditional transfer, since primary school enrolment is very high. The secondary school grant "conditionality" constraint is instead binding for the majority of potential recipients, as secondary school enrolment is much lower.

The effect of aid policies on migration is estimated by using the data collected for the evaluation of Progresa, whose characteristics are discussed below. The main findings in this chapter support the hypothesis that credit constraints limit migration: indeed, unconditional transfers are associated with increased international migration, partly due to new trips being financed or to households substituting (cheaper but with low-yielding) domestic migration for (costlier but with a higher benefit) international one. At the same time, grants to secondary school attendance are associated with a reduction in migration.

1.3 Estimation of treatment on the treated effects using the experimental sample of Progresa

This chapter deals with the estimation of treatment on the treated effects (TTEs) using the Progresa data. Progresa is an aid program implemented in Mexico, aimed at fostering education, health and nutrition among poor households. The data collected for its evaluation consists of 506 villages, 186 of which are randomized out. Poor residents of these villages are not administered the programme until 2000. All households are classified into poor and non-poor according to the information collected in the pre-programme September 1997 census of Progresa localities. All residents of both control and treatment villages are then interviewed at biannual intervals. Detailed data are collected on health, consumption, income and employment, education and migration at least in one of the two annual surveys. With a sample size ranging between 22,000 and 25,000 households in both control and treatment villages, complete coverage of all locality residents, a panel of up to five waves, of which one or two prior to the implementation of the programme, and the exogenous variation induced by the randomization, the Progresa data have attracted the attention of scholars and researchers.

One flaw of this data set is that counterfactuals for a large sub-group of eligible households in treatment communities could not be identified. This has normally resulted into the exclusion of part of the available data, which are simply omitted from the analysis. The contribution of the current exercise consists of recognizing that this group is composed by four different subsets; suggesting ways to identify counterfactuals for three of these subsets; presenting alternative hypotheses to identify TTEs in the presence of non-compliers; providing an empirical application by estimating TTEs on school enrolment for all subsets of households.

The subgroup of households for which a counterfactual could not be identified can be divided into the following categories:

- "True" beneficiaries, i.e. households who are eligible to participate to the program, and receive the treatment, provided that they comply with its requirements.

- Standard non-compliers, that is households who, although eligible, decide not to participate to the program. The standard approach to estimate TTEs in the presence of non-compliers consists of either making some hypotheses on the likely program effect on this group, or finding exclusions restrictions to identify a valid set of counterfactuals in the control group.

- "Forgotten" households who, despite their eligibility to the program, are excluded from it because of administrative errors. These households know they belong to the treatment group but do not receive any subsidies, irrespective of their compliance with the program rules.

- Households who begin being recorded in the surveys after the program begins, hence who are not included in the program although they have the required characteristics to be included in the treatment group (i.e. they are classified as poor).

I discuss the relevance of the estimable parameters, and suggest possible applications.

1.4 Conclusion

This thesis seeks to contribute to the migration literature by assessing the impacts of two existing and popular policies on both international and domestic Mexican migration. It is a first step towards a comparative evaluation of the effectiveness of the policies, as well as an attempt to focus on the indirect effects that aid programs to developing countries may have.

The methodological contribution is very specific, because it regards one particular sample. Nevertheless, it is hopefully going to provide with new parameters that will enable researchers to answer a richer set of questions.

Chapter 2

U.S. Border Enforcement and the Net Inflow of Mexican Illegal Migration

2.1 Introduction

This paper studies the relationship between illegal immigration to the United States and the enforcement of its borders. Border enforcement has been a cornerstone of U.S. immigration policy, especially since the second half of the 1980s. Understanding the impact of border controls on the flow of illegal Mexican migration is of primary importance for several reasons. First, Mexican nationals accounts for 69 percent of the total unauthorized resident population of the United States. Second, most illegal Mexican entries occur through the southern U.S. border. Third, undocumented Mexican migrants tend to be very mobile and to undertake multiple U.S. trips over the life cycle. Donato *et al.* (1992), for instance, show that migrating at least once increases the likelihood of undertaking future migrations.

The undocumented resident population of Mexican nationality has grown from 1.1 millions in 1980 to 2 millions in 1990 and 4.8 in 2000 (U.S. Census, Immigration and Naturalization Service), with an average annual growth of

90,000 units in the 1980s and 280,000 ones in the 1990s. At the same time, the intensity of border enforcement has nearly trebled between the early 1970s and the mid 1990s. The allocated resources to border patrol have been growing steadily especially since the 1986 Immigration Reform and Control Act. In 2002, the immigration enforcement appropriation totalled two billion dollars.

The large growth in undocumented migration despite the higher expenditure in border controls questions the effectiveness of such policy. I argue that tight border controls may have perverse effects on the net flow of illegal migration because they influence the behaviour of both prospective and current migrants: while enforcement increases may deter prospective migrants from crossing the border illegally, the optimal U.S. permanence of current ones may be lengthier because of the higher costs of future migrations. This effect may be of a relevant magnitude, given the large size and the high mobility of such group (see Massey and Singer (1995) for estimates of annual inflow and outflow of U.S.-based Mexican migrants.). If tougher border enforcement lengthens migration duration by raising its cost, patrol of the border might, to some extent, indirectly encourage the formation of a more permanent undocumented resident community.

This fact provides an additional explanation for the disproportionate resources allocated to border versus interior enforcement. If the effectiveness of border patrol is measured by its reduction in undocumented entries only, neglecting its impact on the outflow of migrants, the resulting estimate will overstate border enforcement's true effect on migration net inflow. This causes a larger than optimal resource allocation to border controls.

Although there is awareness of these issues both at the theoretical and anecdotal level, the existing literature has focussed almost entirely on border enforcement's impact of migration inflow only, neglecting to assess its effect on the outflow of illegal migrants. I suspect this is partially due to the scarcity of data on undocumented migrations.

This paper seeks to understand the impact that border policing has on the net inflow of illegal Mexican migrants, contributing to the existing literature in several ways. First, it proves how the impact of border patrol on migration net flow is theoretically ambiguous, and how optimal enforcement is a function of the stock of the U.S. undocumented resident population.

Second, it presents a model where multiple migrations over the life cycle are the optimal choice of utility maximizing rational agents with complete information. To my knowledge, multiple migrations have been explained as the consequence of shocks or of imperfect information. I show that if absence from home entails a positive cost that increases over time, repeated, short-term migrations may maximize inter-temporal utility even in the presence of a sunk cost of migration. This setting permits us to endogenise migration duration.

Third, it provides direct estimates of the effect of border enforcement on the net inflow of undocumented Mexican migrants. This is the first paper in which such estimates are provided. This is achieved by merging border enforcement information with individual-level data on undocumented migration from the Mexican Migration Project (MMP71). Hence, it is possible to test border controls' effects on both the likelihood and the duration of illegal migrations. The obtained estimates are consistent with the theoretical predictions. Border enforcement has a significant deterrent effect, as it discourages prospective migrants from attempting an illegal trip to the United States. At the same time, it lengthens the U.S. permanence of current migrants. My favourite set of estimates reveals that a marginal increase in border controls is associated with a significant reduction in the net inflow of undocumented Mexican migrants between 1972 and 1993. The net effect is roughly half the size of the inflow reduction. Hence, assessing the effectiveness of border enforcement by analysing its deterrent effect only provides a gross overestimates of its true impact

Fourth, it considers how the survey design and the data structure may affect the estimation of the parameter of interest. This analysis, which may be relevant in different applications, has to my knowledge never been undertaken before, despite the wide use of the MMP71 sample.

The paper is organized as follows. Section 2 reviews the existing literature. Section 3 describes the data used. Section 4 derives testable hypotheses on the impact of border enforcement on both inflow and outflow of undocumented migrants using a dynamic model of illegal migration with heterogenous costs. Section 5 illustrates the empirical specification and the related estimation and identification issues. Section 6 presents the results from the econometric analysis. It further discusses how the sample design may bias the estimates of the border controls coefficient. Section 7 uses the estimated parameter of interest and calculates the size of the undocumented illegal migration to compute border patrol's net effect on the size of undocumented migration, highlighting some policy implications. Section 8 concludes.

2.2 Literature review

There is a growing literature trying to understand the relationship between illegal migration and border enforcement. However, the effects of border controls on migration net flow are scarcely known. In my opinion, there are a conceptual and a practical reason for the scarcity of information on the impact of border enforcement.

The main practical problem faced by this literature regards the difficulty of measuring directly the volume of flow and stock of illegal migration.[1] Records of aggregate apprehensions are available. However, only tentative inference can be made to estimate the magnitude of the illegal migration volume from this series, hence the impact of border patrol: in fact, under current U.S. migration law, the same individual may be arrested several times while trying to cross the border. If an apprehended illegal alien agrees on a voluntary departure, that person is simply driven back to the Mexican

[1]Warren and Passel (1987) use U.S. Census data to estimate the stock of undocumented aliens. Massey and Singer (1995) obtain estimates of individual probability of apprehension to assess the magnitude of the net flow of illegal migration.

side of the border. Thus number of arrests does not clearly represents the volume of apprehended migrants.[2] Borjas *et al.* (1991) look at the relation between apprehensions and expenditure for border enforcement. Bean *et al.* (1990) assess the impact of the 1986 Immigration Reform and Control Act. Espenshade (1994) tests the deterrent effect of the probability of border apprehension on the inflow of undocumented migration. Hanson and Spilimbergo (1999), estimate the elasticity of apprehensions with respect to border enforcement from a reduced-form aggregate apprehension function. Davila et al (2001) estimate the short and long-run deterrent effect of border controls.

On the conceptual side, the existing empirical literature has not fully acknowledged the necessity of considering enforcement's impact on migration's net flow, and has rather concentrated solely on the inflow, although there is increasing awareness of the link between migration costs and duration. All the aforementioned papers are concerned with the estimation of the migration inflow effect of border controls. Instead, no mention is made of the fact that rising migration costs may lengthen current illegal migrants' optimal migration duration. This fact is illustrated in Hill's (1987) model of individual migration, where it is shown that higher enforcement may reduce the number of migrations while increasing their length. More recently, Cornelius (2001) mentions the possibility that higher migration costs may result in lower mobility for undocumented individuals, once they have reached the destination country. The same point is made in a recent survey by The Economist (2002).

A notable exception in the empirical literature is constituted by the work of Kossoudji (1992), who uses a sample of repeated illegal Mexican migrants to assess the effect of past apprehension on current migration frequency and

[2]Espenshade (1995b) estimates the relationship between aggregate attempts and apprehensions using a repeated trials model of illegal migration and presents a way of obtaining estimates of the flow of undocumented aliens by observing the fraction of repeated apprehensions of the same individuals. However, collection of this piece of information has been discontinued since the late 1980s.

duration. However, the data set she uses does not permit to distinguish interior from border apprehension. Furthermore, being a sample of sole migrants, it does not permit to estimate the deterrent effect of enforcement (i.e. the size of the individuals discouraged from attempting to migrate by the high level of controls).

A branch of the literature questions the policy effectiveness on different grounds from the one made above. Advocates of interior enforcement suspect that inspections to firms in undocumented labour-intensive sectors might prove more successful in the eradication of illegal migration. The very rationale of border controls is challenged by a work by Hanson and Spilimbergo (2001). Their paper suggests that border enforcement may be the product of conflicting interests, as it is relaxed when the demand for illegal labour is high. Davila et al. (1999) argue that the disproportionate resource allocation favouring border versus interior enforcement is consistent with budget-maximizing behaviour, rather than with trying to minimize the stock of illegal U.S. residents.

As regards migration and individual heterogeneity, theories range from positive migrant self-selection in unobservables because of motivational reasons to negative one, when migrations take place from a country with a wide income distribution (such as Mexico) to a less unequal one (Borjas (1987)). As concerns migration duration, the extremes of the spectrum of existing theories are target earning behaviour, which postulates an inverse relationship between labour market skills and length of stay in the host country, and demand-side selection in skills, according to which unsuccessful migrants are forced to leave before successful ones.

2.3 The data

Table 2.1 describes the data used, providing their means and standard deviations for the period of interest.

The data used come from several different sources: border enforcement and aggregate apprehension data are from unpublished records of the Immigration and Neutralization Service (INS), the Department of Justice agency managing border enforcement. U.S. wage and unemployment data are from the Bureau of Labor Statistics. Mexican unemployment rate is from World Economic Outlook (WEO) data. The other Mexican macroeconomic variables used are from the Mexican National Statistical Institute (INEGI).[3] Hanson and Spilimbergo (1999b) describe both enforcement and macroeconomic data in great detail.

Table 2.1: Descriptive statistics of the variables used, November 1998 data.

	1972-1993		
Name	Mean	St. Dev.	Description
Linewatch hours	2.15	0.51	Million annual hours of border control
Illegal migrations	0.037	0.188	Proportion undertaken illegal migrations
Returns	0.546	0.497	Proportion of returns from illegal migrations
US unemployment	6.84	1.30	U.S. unemployment rate
US wage in pesos	29.53	5.86	U.S. real hourly wage rate in the private sector, 1990 pesos
MX wage index	1.09	0.16	Mexican real hourly manufacturing wage index, production (1990=1)
DEA budget	495.25	241.91	Drugs Enforcement Agency real budget, 1990 million dollars
Legalizations	25.95	43.86	Annual MMP71 Mexican legalizations
Apprehension	540.54	216.09	annual INS border apprehensions (x1000)

In the MMP71 sample, a panel is created from individual retrospective information collected from 71 communities in 13 different Mexican states between 1987 and 1998. Every year a number of different Mexican communities (normally 5) are selected in such a way as to represent a range of diverse characteristics (size, ethnic composition, location and economy). In each of them, a randomly drawn sample of 200 households is interviewed between December and January.[4] Heads of household are asked about their

[3]With the exception of the devaluation rate, which is from aggregate MMP71 data.

[4]In these months migrants tend to return home for Christmas.

migratory experience and their labour history. Interviewees include individuals with past spells of migration (both legal and illegal ones) as well as others who never migrated. Each year, a series of U.S.-based non-random follow-up interviews are undertaken in the summer months, using snowball sampling[5]. This much smaller group is discarded from the analysis because it is selected through a non-probabilistic sampling methodology, although Massey and Zenteno (1999) analyze the quality of the data, and conclude that they are generally a representative source of information on Mexico-U.S. migration. Further details about the study design can be found in Massey (1987). Valid cases are selected by looking at individual age and physically able to migrate: people aged outside the 16 and 55 interval and receiving a disability pension during the recall period are excluded from the valid sample.[6][7] The upper cut-off year is 1993, because the available data for the last five years of the sample (1994 to 1998) consist of too few observations to be included. In such way an unbalanced person-year panel with both entries and exits is obtained.

Since individuals are interviewed in Mexico, I only observe complete spells of migration, and I do not have information on current migrants. The excluded individuals may not be a random sub-set of the illegal migrant population. The issue of the potential non-randomness in the sample and its implications regarding the bias in the estimates of the effect of enforcement on the probability of both migrating illegally and returning from an undocumented migration will be discussed in detail in the next section.

After the described arbitrary truncations, two groups are formed. One is composed by 9,990 individuals and 143,851 person-year units, observed when they are in Mexico until an illegal migration takes place. Observed migrations are recorded by a variable that takes the value of zero whenever

[5]A method whereby interview subjects are indicated by previous interviewees.

[6]The results from the econometric specification are robust to the inclusion or exclusion of women, who are much more unlikely to migrate. Roughly 25 percent of males from the valid sample migrate illegally, versus 6 percent of women.

[7]The maximum age requirement applies only to potential migrants.

the individual is in Mexico, and of one when the individual migrates to the United States. All subsequent years spent in the U.S.A. are recorded as missing. About 85% of the sample never migrates in the observed years, with a maximum of 22 migrations being undertaken. Mean migration for the whole sample is 0.59, while average number of migrations for those who migrate is 2.49.

The other group consists of 2,316 migrants (and 8,986 person-year units) who stay in the United States at least once. In this case, the interest is on the likelihood of returning from a migration. Hence, a variable records with a zero each year spent in the U.S.A., taking instead the value of one in the year the migrant returns home. The variable is missing for all years spent in Mexico between intermediate migrations.

As regards aggregate data, border enforcement's upward sloping trend (measured along the whole U.S. Mexico border) and its higher steepness in the second half of the 1970s and since 1986 can be noticed in Figure 2.1.

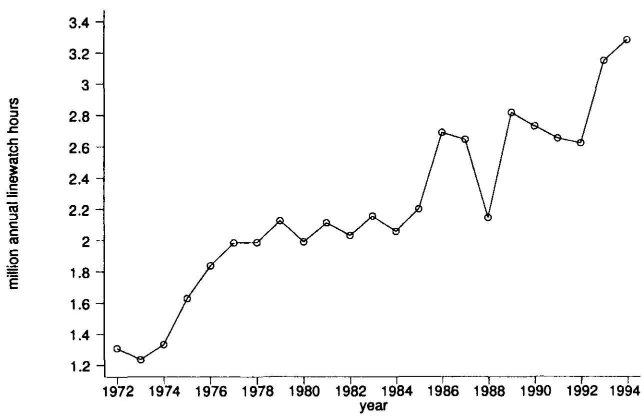

Figure 2.1: Border Patrol - million linewatch hours

Average intensity of border controls between 1986 and 1993 is 20% higher than the 1977-1985 one as a consequence of the implementation of the 1986 Immigration Reform and Control Act (IRCA). Indeed, one of the Act's various measures to curb illegal immigration is the increase in border policing[8],

[8]Border patrol budget, which has been increasing continuously since 1986, does not

together with sanctions for the employment of undocumented labour and with an amnesty to legalize a large group of current illegal U.S. residents. The implementation of IRCA is partly a reaction to the increased migration incentives caused by the economic and financial crisis that hit Mexico in the early 1980s. With high inflation levels eroding the purchasing power of wages[9] and periodic peso devaluations, more individuals are induced to seek jobs illegally in the United States.

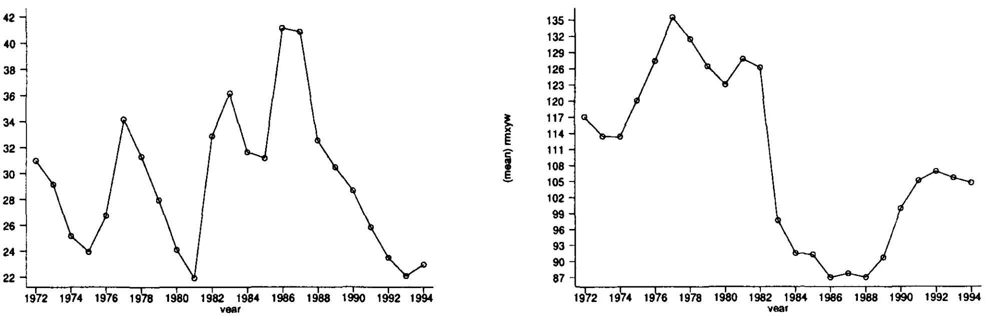

Figure 2.2: U.S. hourly private sector wage in 1990 pesos (left) and Mexican real manufacturing hourly wage index (1990=100)

Figure 2.2 shows the change in the peso value of both U.S. and Mexican wages caused by the high inflation and by the currency devaluations both during and before the Mexican crisis. With higher expected benefits from migration, a higher enforcement level is required to discourage individuals from making an illegal U.S. trip.

---

reflect the 1988 sharp linewatch hours decrease. One possibility is that funds might have been allocated to other enforcement-related activities.

[9]Mexican real minimum wage decreases by 74% in the first half of the 1980s (Hanson and Spilimbergo, 1999).

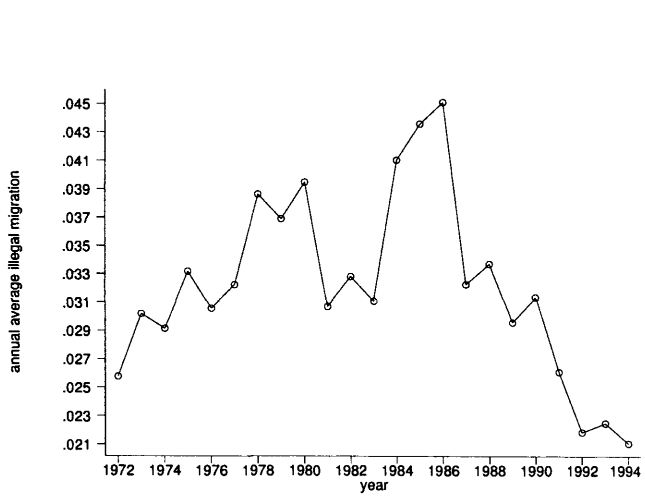

Figure 2.3: Mean annual migration (as proportion of potential migrant population)

The impact of the Mexican economic crisis on incentives to migrate is also noticeable in Figure 2.3, where average annual migration is expressed as proportion of overall population of potential migrants. Although the size and growth rate of the underlying population is unknown[10], the proportion of individuals who migrate nearly doubles between 1972 and 1986, with a peak in 1985-1986, when rumors of the incoming amnesty cause a rush to the United States. The rising migration trend reverses from 1986. Part of this effect is supposedly a consequence of the IRCA amnesty granted to more than two million U.S.-based illegal Mexicans migrants: a large group of individuals simply switch from illegal to legal status.[11] The improvement of the Mexican economic conditions, the toughening of border enforcement and the introduction of employer's sanctions are expected to be further determinants of the trend reversal. Figure 2.4 shows average annual returns from illegal migrations (expressed as proportion of U.S.-resident Mexican

[10]With a positive growth rate of the Mexican population and a deterioration of low-income household living standards during the 1980s, the number of potential illegal migrants is expected to rise, unless outbalanced by legalizations.

[11]INS estimates a total of 2,600,000 Mexican legalizations (Statistical Handbook on U.S. Hispanics, 1992), while Bratsberg (1995) reports nearly 2,200,000 ones.

undocumented population). The pre-IRCA upward sloping part is consistent with the hypothesis of an increase in short-term temporary migration, favoured by lower individual apprehension probability brought by a rising migration volume and a roughly constant border enforcement.

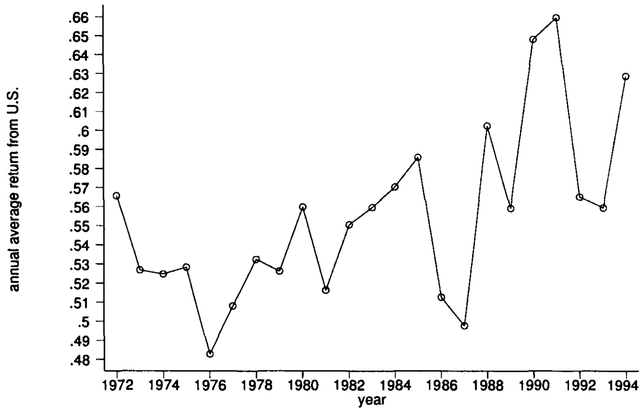

Figure 2.4: Annual returns from illegal migrations (as proportion of U.S.-resident Mexican undocumented population)

The sharp decrease in 1986 and 1987, which coincides with the rise in border patrol activities, is probably exacerbated by the lower mobility of amnesty applicants (as well as by the tighter border controls). In fact, it is likely that amnesty applicants refrain from returning to Mexico while waiting for their legalization application to be processed, i.e. approximately between 1986 and 1988.[12] The last 5 years do not show a clear trend in returns from illegal migrations. Average returns are now proportionally higher than pre-IRCA ones.

Two measures of wages are available: for the U.S.A., average hourly minimum and private sector wages; an index of average production manufacturing hourly wage, or monthly average minimum wage for Mexico. On the one hand minimum wages seem to be more appropriate, since repeated mi-

[12]This result is obtained summing the 12-month interval allowed to apply for the legalization and further few months for the authorization to be granted to the year of IRCA implementation (1986).

grations regard mainly unskilled individuals: average education level for the population of potential migrants is 5.1 years, whereas Hanson and Spilimbergo (1999b) report that mean school years for individuals employed in manufacturing in Mexico was 8.1 in 1990. Moreover, low-skilled, illegal migrants are likely to be offered low wages. On the other hand, though, minimum wages are consistently associated with a higher degree of multicollinearity in the regressions.[13] For this reason, U.S. private-sector wage and Mexican manufacturing wage index are instead used. Their trends are not too dissimilar from their respective minimum wages.

2.4 A model of repeat migration

This section models multiple migration choices of heterogeneous agents. The model is used to understand how the decisions of migrating illegally and of returning from a migration are affected by border controls, and to derive the testable hypotheses and the appropriate econometric specification. In addition, it sheds light on how the survey design may bias the estimates of the parameters of interest.

While it is understood why people migrate (Sjaastad (1962)) and why it may be optimal to return to the home country even given the higher foreign wage (see, for instance, Borjas and Bratsberg (1996), Stark *et al.* (1997) and Dustmann (2003)), less is known about repeated migration. Since moving is costly, it is not clear why individuals may refer multiple, short-termed migrations to a single, longer one. Multiple migrations have been linked by the literature to imperfect information or shocks (Da Vanzo (1983)). In my view, however, repeated migrations over the life cycle can arise also as the optimal decision of agents with complete information. Given the vast scale of Mexican migration to the United States and the existence of migration networks, it is believed that access to information is relatively easy and

[13]Analysis of the correlations between both sets of wages and the other variables used in the econometric model show that the multicollinearity arising from the use of minimum wages cannot be directly related to any pairwise correlation.

costless.

Two main determinants of repeated Mexican migrations over the life cycle are: labour demand and costly absence. The former explains the existence of cyclical migration patterns as the consequence of seasonal changes in labour demand. One obvious example is the increased demand for agricultural workers during harvest time, or the one for hotel staff during peak season. The latter, instead, is based upon the concept that being away from home has a positive cost. This can be thought of as absentees' positive likelihood of losing claims on current and future ownership of family assets (such as land and properties) increasing with time spent abroad, or as individuals being homesick. The longer the absence, the looser the tie with the family members left behind. Such hypothesis is consistent with the alternative views of remittances being used as a means to retain a tie with the household in the home country, or being a manifestation of altruism. The hypothesis that time away from home may be costly is consistent with the observed evidence from the MMP71 data: the likelihood of returning from a migration is *ceteris paribus* higher for parents and married migrants, and is increasing with the number of children and hectares of land owned.

The current model of repeated migration is based on the latter hypothesis. While further research on the issue is clearly needed, I believe that the model sketched here captures some existing features of Mexican repeat migration. In any case, the implications of the model regarding the effect of border enforcement on the intensity and duration of illegal migration could have been equally derived from a wide class of alternative determinants of multiple migrations.

An additional feature of this model is that it endogenises migration duration. Hill (1987) models multiple migration over the life cycle assuming that migration duration if fixed.

I assume that migration costs are heterogeneous. This may arise because of different reasons. Individuals may be endowed with different levels of ability to cross the border, or they may have access to private information. The data provide evidence of heterogeneity in border-crossing ability. The number of observed apprehensions within a single trip in fact varies considerably, ranging between 0 and 15, although 78 percent of migrants manage to cross the border at first attempt, and 16 percent with up to three. Alternatively, individuals may face varying degrees of financial constraint, hence financing similar trips may entail different costs. Finally, costs may vary because of distance from the border and different levels of community-specific network effects. Modelling alternative sources of heterogeneity, such as differences in labour market-related skills, results into different types of individuals being selected in and out of migration. Nevertheless, the impact of border controls on their net inflow in the U.S. would be the same irrespective of the type of heterogeneity (in sign, if not in magnitude).

Assume a continuum of potential illegal migrants with heterogeneous migration costs deriving from different ability endowments. Agents treat wages and border enforcement (bp) as given. There is no uncertainty. Assume the existence of a utility function representing individual preferences. Such function depends on consumption ($c > 0$) and on the fraction of time spent in the host country (t) and it is assumed to be continuously twice differentiable, additively decomposable both in consumption and time and over time. Individuals choose the optimal migration duration in both periods given the positive wage differential between U.S. and Mexican wages ($w^{US} > w^{MX}$), the strictly concave disutility from staying away from home (u_t[14] < 0 and $v_{tt} < 0$), their appetite for consumption ($u_c > 0$ and $u_{cc} < 0$)[15], and the sunk cost associated with the illegal migration (M_C). The latter is a positive function of border enforcement (bp), $M_{Cbp} > 0$, and a negative one ($M_{Ca} < 0$) of an individual-specific parameter ($a \in (0,1)$) capturing heterogeneity in costs. a varies continuously among the agents. This parameter will be from now on referred to as ability in crossing the border. I assume capital market perfection. r is the interest rate paid on savings (S) and β

[14] Partial derivatives are sub-indexed with respect to the argument, hence $\partial v/\partial t = v_t$.

[15] The assumptions on the utility function ensure that the agents' optimization problem has a maximum.

is the inter-temporal discount factor.

The maximization problem is:

$$\max_{c_1,c_2,S,t_1\in[0,1],t_2\in[0,1]} U(c_1,c_2,t_1,t_2,S;a) = u(c_1(t_1,S,a),t_1)+\beta u(c_2(t_2,S,a),t_2)$$

where

$$\begin{aligned}
c_1 &= t_1 w_1^{US} + (1-t_1)w_1^{MX} - M_C(bp_1,a)_{1\{t_1>0\}} - S \\
c_2 &= t_2 w_2^{US} + (1-t_2)w_2^{MX} - M_C(bp_2,a)_{1\{t_2>0;t_1<1\}} + (1+r)S \\
w^{US} &> w^{MX}
\end{aligned}$$

Given the parameters of the model, different ability endowments will be associated with varying optimal migration durations, including $t_1^* = t_2^* = 0$ and permanent migrations. In the current analysis, however, I restrict the valid solutions to include only return migrants, i.e. $t_1^* \in [0,1)$ and $t_2^* \in [0,1)$.

The first-order conditions[16] for migrants in both periods, i.e. individuals for whom $t_1^* > 0$ and $t_2^* > 0$, are:

$$\begin{aligned}
\frac{\partial U}{\partial t_1} &= u_{t_1} + u_{c_1}(w_1^{US} - w_1^{MX}) = 0 \\
\frac{\partial U}{\partial t_2} &= \beta u_{t_2} + \beta u_{c_2}(w_2^{US} - w_2^{MX}) = 0 \\
\frac{\partial U}{\partial S} &= u_{c_1} - \beta(1+r)u_{c_2} = 0
\end{aligned}$$

The model could be easily extended to allow for future migration costs being a negative function of past migrations or for agents being unable to borrow. The conclusions reached would be unaffected, although the magnitude of the migrant flows would differ.

2.4.1 Deterrent effect

Given preferences, wages, some continuous ability distribution and enforcement, agent with different ability levels will choose a different combination

[16]The second-order condition matrix is negative semi-definite.

of current and future migrations. Individuals whose ability level exceeds a certain threshold ($a > a^h$) will migrate in both periods. Those with a sufficiently low ability ($a < a^l$) will never migrate. The impact of changes in migration costs on the inflow of migration can be studied by observing how the ability threshold levels change due to higher enforcement. If I prove that the values of the existing thresholds are augmented by a marginal increase in border controls, then I have shown that higher enforcement is associated with a lower migration inflow.

Agents' behaviour cannot be predicted without assuming an explicit form of the utility function or without making restrictions on the range of values taken by the parameters of the model. Indeed, the model is such that agents choose in each period whether to migrate or not, and for how long. Using H and M to indicate home stay and migration, there will be four possible outcomes: HH, HM, MH, MM and various ability threshold for individuals indifferent between alternative inter-temporal choices. There are six potential thresholds (the six possible pairwise combinations of the four different outcomes). Only a sub-set of them will be valid, depending on the values of the parameters and on the type of utility function.

I may illustrate the impact of higher migration costs on the ability thresholds by performing comparative statics exercises for all possible cases. I believe that this is unnecessary, given that higher migration costs will in no circumstance result in higher aggregate migration. Since the effect of tighter enforcement on immigration is not ambiguous, I prefer to choose one specific case to illustrate this point formally.

I consider the case in which the values of wages and enforcement are identical in both periods and $\beta(1+r) = 1$. In such case, optimal migration duration will be the same in both periods ($t_1^* = t_2^*$) and there will be no savings. Since the two periods are identical, agents will either migrate optimally in both periods or stay always in Mexico. Hence, there is a unique ability threshold (i.e. $a^h = a^l \equiv a^T$). Consider the agent indifferent between migrating in both periods and not migrating in either

$$MM(a^T) - HH(a^T) = 0 \tag{2.1}$$

In this scenario, any current or future increases in border controls will induce the agent to favour the zero migration outcome. Formally, I can compute the sign of the enforcement effect on ability threshold (a^T) from (2.1) using the implicit function and the envelope theorems:

$$\begin{aligned}
\frac{\partial a^T}{\partial bp_1} &= sign\left(-\frac{u_{c_1^*}(\partial Mc/\partial bp_1)}{\left(u_{c_1^*} + \beta u_{c_2^*}\right)(\partial Mc/\partial a)}\right) > 0 \\
\frac{\partial a^T}{\partial bp_2} &= sign\left(-\frac{\beta u_{c_2^*}(\partial Mc/\partial bp_2)}{\left(u_{c_1^*} + \beta u_{c_2^*}\right)(\partial Mc/\partial a)}\right) > 0
\end{aligned}$$

The effect of a current enforcement increase is larger than the one of a future cost rise because of discounting. Permanent enforcement increases will result in a more marked rise in the ability threshold.

While the simple case presented above is sufficient to make my point, more realistic assumptions on variation in costs and benefit of migration and in inter-temporal preferences will result in agents with varying ability endowments being indifferent between alternative options. Such alternatives may imply different responses to current or future enforcement changes. All possible cases show that higher enforcement is associated with positive (or at most non-negative) changes in the minimum level of ability required to migrate.

One case worth mentioning is when the anticipation of higher future costs of migration results in migrations being undertaken in period one rather than in period two. Consider the individual who is indifferent between migrating once, irrespective of when the migration occurs: $MH - HM = 0$. This individual will be induced to migrate in the first period by a marginal increase in future migration costs. The reverse is true in case of an increase in current enforcement levels.

2.4.2 Optimal migration duration

Unlike the migration decision, optimal migration duration is a function of both current and future enforcement level. While higher levels of border controls deter some prospective migrants to attempt to reach the United States, reducing the inflow of illegal migration, the average migration duration of current migrants will increase, given the higher migration costs. Comparative statics shows that the optimal migration length is a positive function of both current and future levels of border enforcement:

$$\frac{\partial t_1^*}{\partial Mc_1} = sign\left(-(1+r)u_{c_2^*c_2^*}(u_{c_1^*c_1^*}u_{t_2^*t_2^*} - 2u_{c_2^*c_2^*}\Delta^2)\right) > 0 \quad (2.2)$$

$$\frac{\partial t_1^*}{\partial Mc_2} = sign\left(\beta u_{c_1^*c_1^*}u_{t_2^*t_2^*} + u^2_{c_2^*c_2^*}\Delta^2(2+r)\right) > 0 \quad (2.3)$$

where Δ indicates the wage differential.[17] These results are intuitive: the higher the costs of migrating, the longer the time that must be spent abroad to reap positive returns from the migration. Changes in border enforcement influence the behaviour of current illegal migrants as well. Indeed, migration duration is also a positive function of future migration costs. Comparative statics shows that for individuals who migrate in both periods, a marginal increase in future enforcement levels will increase optimal current migration duration. This means that, while prospective migrants may be deterred from migrating by an increase in border patrol, current ones will lengthen the duration of their present migration spell. In practice, given the large stock of undocumented resident U.S. migrants, the latter effect may be of a relevant size.

Equations (2.2) and (2.3) are particularly interesting, given the objective of the current exercise, because they show how increases in border policing, although reducing the inflow of undocumented migration, extend the length of permanence in the United States for those who decide to migrate and for current illegal migrants.

[17] I also assumed that wage differentials are identical in the two period. This simplifies the notation without changing the signs of the partial derivatives.

To summarize, modelling the effect of border enforcement on choice and duration of repeated temporary undocumented migrations shows how the former has a deterrent effect: higher enforcement levels correspond to fewer migrations. However, rising migration costs affect also the length of the permanence in the receiving country through an increase in both current costs, which require migrants to stay longer before they can benefit from the investment, and in future ones, which induce present migrants to delay their return in view of the higher difficulty of the next illegal trip.

The effect of an increase in border policing on the net inflow of illegal migration is ambiguous and depends on the size of both prospective and current migrants and on how sensitive both groups are to changes in migration costs.

Some implications of these findings are: first, policy-induced changes in illegal migration inflow differ from changes in its stock, because the policy has also a significant effect on migration outflow. Second, the estimated effect of the policy observing only migration inflow is always larger than the true one (in absolute value). Third, considering migration inflow only results in a sub-optimal resource allocation, where too much spending is devoted to border enforcement and too little to alternative policies. Fourth, for each level of prospective migrants, the optimal level of border enforcement is a function of the stock of current U.S.-based illegal migrants.

2.5 Empirical specification and estimation issues

The model has clearly highlighted the ambiguous effect of border enforcement on the net inflow of illegal migration. I now proceed to estimate the observed magnitude of such effect for the 1972-1993 time period. My empirical approach consists of estimating the effect that changes in border controls have on the individual probability of transition between different states: from staying in Mexico ($m = 0$) to becoming an illegal migrant ($m = 1$); and from being an undocumented U.S. resident ($r = 0$) to returning home ($r = 1$). Time is discrete and the unit is the year. I model the two transitions separately. i indicates the i-th individual and t the year.

2.5.1 Estimating the likelihood of migrating illegally

The previous section has shown two facts that will be used to build the empirical specification. First, migration decision is a function of current and future values of the parameters. Second, an increase in current border enforcement has a direct negative effect on current illegal migrations and an indirect positive effect through its impact on expectations of future migration costs. I use a latent regression model linear in parameters to specify the conditional likelihood of migrating at time t, given that the individual is in Mexico at the end of the previous period, $P(m_{it} = 1|m_{it-1} = 0) = P(m^*_{it} > 0)$, as a function of the parameters of the model:

$$m^*_{it} = \tau^m Mc_{it} + \gamma^m X_t + \sum_{s=t+1}^{T} \left[\delta^m_s E\left(Mc_{is}\right) + \pi^m_s E\left(X_s\right)\right] + f^m_i + h^m_t + \varepsilon^m_{it}$$

where Mc indicates migration costs, as before, and X contains macroeconomic differentials that influence the migration decision and are potentially correlated with border controls. T indicates the number of future periods' parameter values that affect the current migration decision. f and λ represent all individual- and time-specific factors that influence the migration decision and are not captured in the theoretical model. ε is some white-noise parameter. If I further assume that migration costs can be expressed as a linear function of enforcement and of an individual-specific parameter,

$$Mc_{it} = \theta bp_t + \phi a_i \tag{2.4}$$

I can rewrite the latent model as

$$m^*_{it} = \beta^m bp_t + \gamma^m X_t + \sum_{s=t+1}^{T} \left[\kappa^m_s E\left(bp_s\right) + \pi^m_s E\left(X_s\right)\right] + \mu^m_i + \lambda^m_t + \varepsilon^m_{it} \tag{2.5}$$

where $\beta^m = \tau^m\theta$, $\kappa^m_s = \delta^m_s\theta$, and $\mu^m_i = f^m_i + \tau^m\phi a_i + \sum_{s=t+1}^{T}\delta^m_s\phi a_i$. The effect of a change in the current enforcement level on the migration likelihood is the combined effect of the former's direct and indirect effect, via changes in expected future migration costs,

$$\frac{\partial P(m_{it}=1|m_{it-1}=0)}{\partial bp_t}=\beta^m+\sum_{s=t+1}^{T}\kappa_s^m\frac{\partial E\left(bp_s\right)}{\partial bp_t} \tag{2.6}$$

considering a linear probability model for simplicity. I expect the first term β to be negative, and the following ones positive, although the net effect should be again negative.

A static model estimates (2.6) directly, without the need to make explicit assumptions on how agents form expectations. Its disadvantage is that it does not permit to assess the impact of (expected) future enforcement levels on current decisions. I will estimate both a static model and one where I condition explicitly on future levels of border controls. The object $\frac{\partial E(bp_s)}{\partial bp_t}$ in this latter case is estimated as the coefficient of a regression of future on current enforcement level, conditioning on all the other macroeconomic variables. Given that the parameter of interest is identified only through time variation, I assume that $T=t+1$.

Lagged values of enforcement and of macroeconomic variables are also added to (2.5) in order to control for potential omitted variable bias. A higher past enforcement level, for instance, might induce the agents to postpone the migration decision and to increase savings to finance a future trip.

Current and future border control levels may be endogenous, as the enforcement level and expectations may be correlated to unobservable shocks to migration. This point was first made by Hanson and Spilimbergo (1999b). In order to control and test for endogeneity, Drugs Enforcement Administration (DEA) budget is used. Drugs are smuggled in massive quantities through the U.S. Southern border, and one of the aims of border patrol is also to curb narcotics trafficking.

Border linewatch hours and DEA budget show a high and significant positive correlation (85%) and the coefficient of DEA budget in the first-stage equation is always positive and strongly significant. When both current and future border enforcement are added in (2.5), I use the smallest number of necessary lags of the instrumental variable to identify the parameter of interest. There is a clear trade-off between adding further leads of linewatch hours and macro variables and having to use additional lags of DEA budget and further conditioning variables in the first-stage equation, especially considering that both measures of border enforcement and of the instrument vary only over time. A large number of instrumenting variables may prevent the separation of the endogenous from the exogenous component of border controls.

Attempts to use alternative instruments were made. Dummies to control for the political party of the outgoing U.S. president, and a variable that indicates whether both Senate majority and U.S. president are Republican were both included. The significance of these latter variables suggests that border enforcement resource allocation may be influenced by the political cycle.[18] However, it is not clear whether the former variables do not affect migrants' propensity to move to the United States. Experiments were also made using the number of years to a presidential and congressional election as additional instruments. However, the significance of these further regressors is not very high in the first-stage regression.[19] The result section reports only output obtained using DEA budget as instrumental variable.

An additional issue is posed by the selection of interviewees caused by the sample design. As explained above, there is no information on migrants when the whole household moves. If migrants are selected in unobservables in some way related to border enforcement, this will result in a correlation between the mean of μ_i^m and bp. However, in practice one expects the magnitude of the bias to be small, if not negligible. This occurs for a number of reasons. First, the bulk of Mexican migration appears to be constituted by temporary moves. Individuals tend to migrate on their own, leaving their

[18]There is a significant higher enforcement level when both majority senate and president are Republican. Moreover, patrol of the border is lower in presidential election years, when the outgoing president is Republican.

[19]When I use monthly levels of border enforcement rather than annual one, the significance level of year-to-election variables is much higher.

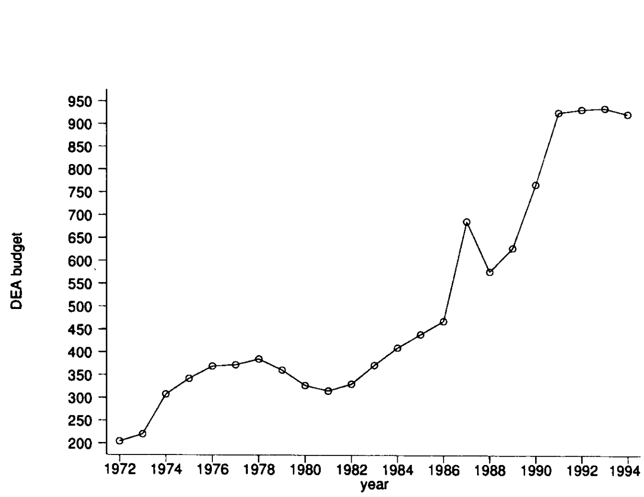

Figure 2.5: Drugs Enforcement Administration budget, 1990 million USD

family behind. Second, as enforcement tightens, both out-migrations and returns will decrease, partly offsetting each other. Third, actual migrations are only a small fraction of the overall sample.

Shocks to migration may be serially correlated. Not only the impact (or the information) may reach the farthest regions with a lag, but also individual reaction to it may be delayed. Individuals geographically closer to the border will react before those in more distant states or communities (distance may also be related to access to information). Furthermore, such shocks may exhibit some degree of persistence. Serial correlation is tested and controlled for appropriately according to the type of estimator used.[20]

The macroeconomic regressors included (X) are U.S. private sector hourly real wage rate in pesos and an index of Mexican manufacturing real hourly wage, together with U.S. unemployment, to capture the probability of finding a job once in the United States. Number of obtained legalization is added as well (l). Since legalizations and enforcement increase at the same time in the second half of the 1980s, as they are both IRCA provisions, it is important to disentangle these two effects. The higher the legalization

[20] I also experimented with more general specifications of the within-panel correlation structure, which turns out not to be perfectly represented by an AR(1) process. These results are not presented but are available upon request.

number, the lower the volume of both actual and potential illegal migrants. Thus, part of the post-IRCA decline in both undocumented border crossings and returns from illegal migration is simply the result of a status change, i.e. becoming a legal resident[21], rather than the direct consequence of tougher enforcement. This effect may be substantial, since more than 2 million Mexicans are estimated to have benefited of the IRCA legalization program.

I estimate the individual likelihood of migrating illegally at time t, conditional of being a potential migrant in $t-1$, using alternatively the linear (OLS and within group) and non-linear (probit) probability models, relying on different parametric assumptions on the distribution of the error term. The former is computationally less demanding for what concerns the estimation of endogenous regressors. However, it has some well-known shortcomings, such as constant marginal effects, predicted probability not constrained to vary between the $[0, 1]$ interval, and heteroskedastic errors. Robust standard errors are computed (also for the probit estimates) and the observations are clustered at the individual level. The estimated coefficients are interpreted as local linear marginal effects.

The alternative estimator relaxes the assumption that the conditional probability is linear in the parameters. I address the endogeneity issue in a non-linear framework by adopting Blundell and Smith's (1986, 1989) value function approach and I compute block-bootstrap estimates of the coefficient standard errors of both border controls and residuals from the first-stage regressions.

[21] I used alternatively both applications and obtained legalization. The former is assumed to proxy changed incentives to mobility, with applicants remaining in the U.S.A. until their request is processed, while the latter accounts for status change. Since legalizations always perform better than applications, I interpret this result as evidence of the higher importance of changes in the pool of illegal migrants than incentive changes, probably in part because the latter's effect are only concentrated in two years.

2.5.2 Estimating the likelihood of returning from an illegal migration

The empirical specification contains the main two features of the theoretical model. These are: first, optimal migration duration is a function of current and future values of the parameters of the model. Second, changes in current enforcement levels affect the length of the current migration spell by changing individual expectation of future migration costs. I express the relationship between migration costs and duration in terms of the individual likelihood of returning home. Hence if migrations are lengthened by higher costs, I expect the return probability to be a negative function of the latter.

For individual i in time period t, the probability of returning from an illegal U.S. trip, conditional on being an undocumented U.S. resident at the end of the previous period, is described by the following latent regression model

$$\begin{aligned} r_{it}^{*} &= \delta^{r} Mc_{it} + \pi^{r} X_{t} + \sum_{s=t+1}^{T} \left[\delta_{s}^{r} E\left(Mc_{is}\right) + \pi_{s}^{r} E\left(X_{s}\right)\right] + f_{i}^{r} + \lambda_{t}^{r} + \varepsilon_{it}^{r} \\ &= \beta^{r} bp_{t} + \pi^{r} X_{t} + \sum_{s=t+1}^{T} \left[\kappa_{s}^{r} E\left(bp_{s}\right) + \pi_{s}^{m} E\left(X_{s}\right)\right] + \psi_{i}^{r} + \lambda_{t}^{r} + \varepsilon_{it}^{r} \quad (2.7) \end{aligned}$$

where $T \geq t+1$ and $\beta^{r} = \delta^{r}\theta$. The second equality uses (2.4). $\kappa_{s}^{r} = \delta_{s}^{r}\theta$ and $\psi_{i}^{r} = f_{i}^{r} + \sum_{s=t}^{T} \delta_{s}^{r} \phi a_{i}$. The parameter of interest is the sum of the direct and indirect effect of changes in border enforcement on the likelihood of returning to Mexico.[22][23]

$$\frac{\partial P(r_{it}=1|r_{it-1}=0)}{\partial bp_{t}} = \beta^{r} + \sum_{s=t+1}^{T} \kappa_{s}^{r} \frac{\partial E\left(bp_{s}\right)}{\partial bp_{t}} \quad (2.8)$$

Both β^{r} and the κ_{s}^{r} are expected to be negative. Again, (2.7) is estimated both as a static equation, and adding one lead and one lag of the

[22] The parameter β^{r} captures both the effect of enforcement on those who have migrated in the same period, and any indirect effect it may have on the return decision on those who had migrated before. It turns out from the empirical analysis that this latter effect is null.

[23] Again, the next formula uses the linear probability model for simplicity.

enforcement and macroeconomic variables. The advantages and disadvantages of both specifications have been discussed in the previous section. All the econometric and identification issues discussed above are valid for the return equation, and are tackled analogously. In addition, in this case the correlation between unobservable individual heterogeneity and (current and expected) linewatch hours might seriously bias the estimate of the parameters of interest. Both the decision to migrate and the one to leave depend on the intensity of border controls. Hence, current migrants' unobservables are correlated with the enforcement level. This issue is dealt with by using within-group and conditional logit estimators. The latter estimator reduces considerably the valid sample size.

Alternatively, I can re-write the latent regression model as

$$r_{it}^{*} = \beta^r bp_t + \pi^r X_t + \sum_{j=t_0}^{t} (\eta_j^r y_j) + \sum_{l=1}^{k} \zeta_l d_l + \sum_{s=t+1}^{T} \left[\kappa_s^r E(bp_s) + \pi_s^m E(X_s)\right] + \mu_i^r + \lambda_t^r + \varepsilon_{it}^r \tag{2.9}$$

y_j are year-of-arrival dummies (for migrations occurred between t_0 and t) and capture the average unobservables of individuals who migrated in the same year. The dummies d_l group individuals who have spent l years in the same spell of migration at time t. I believe that controlling for year of arrival and migration duration reduces the correlation between the enforcement variables and the residual unobservable heterogeneity. The advantage of using (2.9) is that I can indirectly observe the effect of enforcement on selection on unobservables by adding the y_j and d_l dummies sequentially and observing how the enforcement coefficient changes.

The other regressors are those described in the previous section. As before, it is important to condition on legalization number to control for the reduced mobility of amnesty applicants and the status change of successful ones (both resulting in fewer returns to Mexico contemporaneously with the post-IRCA border enforcement increase).

2.6 Results

2.6.1 Probability of migrating illegally

I first estimate the likelihood of undertaking an undocumented migration using the linear probability model. Border control levels measured in million linewatch hours. Table 2.2 reports the coefficients of the static model from different estimators. The Durbin-Watson test provides evidence of positive serial correlation. Even columns treat enforcement as endogenous, and report the results from IV estimation.

The results are in line with the prediction of a negative effect of current enforcement on the likelihood of migrating illegally. The significance of the Durbin-Wu-Hausman test provides evidence of the inconsistency of the linewatch coefficients, when the endogeneity of enforcement is not taken into account. The results suggest that there is a positive correlation between border patrol and aggregate shocks, consistent with the conjecture that the level of enforcement is set according to the expected intensity of the migration flow.

The coefficients from OLS and within-group estimators do not differ (columns 2 and 4). This fact is interpreted as evidence that the sample design does not bias the parameter of interest.[24] The values of the parameters of interests are comparable in sign and magnitude to the ones obtained clustering observations by year. I further tested for the result robustness by changing the set of conditioning variables and the available years. For instance, I included Mexican unemployment rate, which is available only since 1980, and further controls (state of origin, capturing travel expenses to reach the border, dummies for size of community of origin to measure the dimension of the local labour market, and proportion of male labour force in agriculture to control for varying incentives to migrate to differentiate the source of labour income).[25] The estimated coefficients are stable across the

[24]The coefficient from pooled probit is smaller, but this may be due to the different sample size, distributional assumptions or approach to deal with endogeneity.

[25]Results not shown but available upon request.

various specifications and data sub-sets.

The coefficients of border enforcement from the lead and lag specification are generally in line with expectations, although the results are less robust. I suspect that this is partially due to the fact that longer lags of the instrumental variable are used in the first-stage regressions. The very high R-squared values suggest that the first-stage may not appropriately separate the endogenous from the exogenous variation of enforcement. Nevertheless, the results from Table 2.3 remain interesting insofar as they confirm that agents are sensitive to expected increases in future enforcement, which induce them to anticipate the migration. This is shown by the positive sign of the coefficient of future border enforcement levels (although this is not the case in the within-group estimates).

Table 2.2: Probability of illegal migration, static model

	OLS-AR(1)	IV-AR(1)	WG-AR(1)	IVWG-AR(1)	Probit	IV-Probit
	(1)	(2)	(3)	(4)	(5)	(6)
Linewatch hours (bp)	-.0009	-.005***	-.002*	-.005***	-.002**	-.009***
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
Observations	136159		133861		143871	
Durbin-Wu-Hausman	[.000]		[.000]		[.002]	
First-stage R^2			.73			
1^{st}-stage IV significance			[.0000]			

The estimated regression is the one described in the previous section. Robust standard errors in brackets. *= significant at 90% confidence level; ** = significant at 95% confidence level; *** = significant at 99% confidence level.

Table 2.3: Probability of illegal migration, lag and lead of bp

	OLS-AR(1)	IV-AR(1)	WG-AR(1)	IVWG-AR(1)	Probit	IV-Probit
	(1)	(2)	(3)	(4)	(5)	(6)
Linewatch hours (bp)						
t-1	.002	-.012**	.001	-.014***	.005	-.014**
	(.002)	(.005)	(.002)	(.005)	(.002)	(.006)
t	-.009***	-.014***	-.004**	.004	-.012**	-.018***
	(.002)	(.002)	(.002)	(.003)	(.003)	(.003)
t+1	.008*	.022***	-.007***	-.005	.006**	.024***
	(.002)	(.004)	(.002)	(.004)	(.002)	(.005)
H_0: $\|bp_t\| \neq \|bp_{t+1}\|$		[.115]		[.461]		[.349]
Observations	136159		124426		137577	
Durbin-Wu-Hausman	[.000]		[.000]		[.002]	
First-stage R^2						
t-1			.94			
t			.92			
t+1			.94			
1^{st}-stage IV significance			[.0000]			

Robust standard errors in brackets where applicable. Block-bootstrap standard errors in (5) and (6). *, **, ***= significant at 90%, 95%, 99% confidence level.

I interpret the negative sign of the coefficient of lagged linewatch hours as suggestive of slow response to enforcement, both because of the lag between the acquisition and the response to new information, and because tighter enforcement translates into longer waits before a successful border crossing.

Table 2.4 provides estimates of the parameter of interest under the different specifications and estimators used (left panel) and presents the output of a time series regression of border patrol on its lagged values[26] (right panel). These are built following (2.6), using the coefficients from Tables 2 and 3 and the coefficient of lagged border patrol from the two-lag regression below (i.e. $\frac{\partial E(bp_{t+k})}{\partial bp_t} = (0.561)^k$).

Note how the parameters estimated using the linear probability model and the probit one are similar in size.

The estimated parameters of interest suggest that a one million increase in annual linewatch hours reduces the likelihood of undertaking an undocumented migration by 0.5 to 0.9 percentage points, considering the results from the static model. Given that the average annual migration volume amounts to 3.7 percent of the Mexican adult economically active population with no US green card, such linewatch hours increase reduces average Mexican illegal migration by 13 to 25%. Since mean enforcement is 2.15 million linewatch hours, the implied average enforcement elasticity of deterrence is negative but larger than minus one.

2.6.2 Probability of returning from an illegal migration

Table 2.5 shows within group, conditional logit and pooled probit estimates of the enforcement coefficient on the likelihood of returning from and undocumented U.S. trip. As in the previous case, I specify the linear probability model residuals as represented by an AR(1) process. The estimated coefficients are in line with expectations, suggesting that tighter border patrol reduces the likelihood of returning from an illegal migration. The Durbin-Wu-

[26]The results are roughly unchanged after conditioning on U.S. wages and unemployment.

Table 2.4: Effect of a marginal increase in current enforcement on the likelihood of migrating illegally

Model	Static	Lag and lead of bp	bp	
$\frac{\partial P(m_{it}=1\|m_{it-1}=0)}{\partial bp_t}$				
			t-1	.561**
IV AR(1)	-.005***	-.002*		(.217)
	(.001)	(.001)	t-2	.351
				(.249)
Probit	-.009***	-.004*		
	(.001)	(.003)		

Block-bootstrap standard errors. Intercept included in the time series in the right panel, as well as controls for macroeconomic variables. *, **, ***= significant at 90%, 95%, 99% confidence level.

Hausman test provides evidence of the consistency of the IV coefficients. The enforcement coefficients from the non-instrumented regressions are upward-biased. This may be due to positive shocks to migration inducing the policy maker to increase border controls.

Table 2.5: Probability of returning from an illegal migration, static model

	WG-AR(1)	IVWG-AR(1)	CLogit	IVCLogit	IVProbit		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Linewatch hours	-.086***	-.106***	-.174***	-.199***	-.054	-.437***	-.165***
(bp)	(.029)	(.031)	(.001)	(.001)	(.036)	(.071)	(.042)
Observations	6816			5246			8911
Durbin-Wu-Hausman		[.070]		[.046]	[.670]	[.087]	[.750]
Year-of arrival					No	+	+
Duration dummies					No	No	-
First-stage R^2		.76		.76	.93	.95	.94
1^{st}-stage IV signif.		[.0000]		[.0000]	[.0000]	[.0000]	[.0000]

Robust standard errors in brackets where applicable. Block-bootstrap standard errors in (3) to (7). *, **, ***= significant at 90%, 95%, 99% confidence level.

The last three columns start from the basic static specification and add year-of-arrival and migration duration dummies sequentially. The base categories are the "oldest" migration year and length of stay of less than one year. The signs of the coefficients are indicated at the bottom of the table, and show that more recent cohorts are more likely to return from an illegal trip, while longer spells of migration reduce the likelihood of returning. The reduction in the magnitude of the enforcement coefficient and its later increase when adding first the year-of-migration and then the duration dummies suggests that migrants are positively selected in unobservables by the tighter border controls: as controls intensify, "better" individuals manage to migrate. At the same time, higher enforcement is positively correlated with migration duration.

However, the Durbin-Wu-Hausman test cannot be rejected. This may depend on the high first-stage regression R-squared, caused by the additional regressors included in the specification (year-of-arrival and migration duration dummies). The same *caveat* as before applies to the interpretation of the results when the values of the R-squared of the first-stage regressions are very close to one. In any case, once the two sets of dummies are added, the enforcement coefficient is in line with the magnitude of those obtained from within group and conditional logit estimators.

Also in this case, enforcement coefficients in the various specifications are not sensitive to changes in the set of conditioning variables[27] or to the interview years used.

The coefficients from the lead and lag specification are consistent with expectations. Both current end future enforcement levels reduce the likelihood of returning from an illegal migration, as shown in Table 2.6. The results are not as robust as in the static version, though. The positive sign of lagged linewatch hours suggests that high past enforcement may increase

[27] I experimented adding different subsets of the following variables: Mexican unemployment (later discarded because of dubious validity), mean age at migration, mean education, and characteristics of the community of origin (size, proportion of labour force employed in agriculture, state dummies).

Table 2.6: Probability of returning from an illegal migration, lag and lead of bp model

	WG-AR(1)	IVWG-AR(1)	CLogit	IVCLogit
	(1)	(2)	(3)	(4)
Linewatch hours (bp)				
t-1	.072**	1.349***	.072	-.346
	(.037)	(.525)	(.079)	(1.242)
t	-.093***	-.131***	-.169***	-.136*
	(.030)	(.042)	(.060)	(.081)
t+1	.008*	-.685***	-.066	-.203
	(.005)	(.262)	(.049)	(.652)
Observations	6816		5246	
Durbin-Wu-Hausman	[.009]		[.736]	
First-stage R^2				
t-1		.93		
t		.92		
t+1		.94		
1^{st}-stage IV significance		[.0000]		

Robust standard errors in brackets where applicable. Block-bootstrap standard errors in (3) and (4). *, **, ***= significant at 90%, 95%, 99% confidence level.

the return from a trip through higher savings.

The coefficients obtained from the different estimators and specifications vary in size, although not in significance, probably due to the smaller sample size. In Table 2.7 I use (2.8) and the obtained coefficients to compute estimates of the overall marginal effect of border controls.

My favorite set of estimates are the ones from the static model, given the higher variability of the coefficients from Table 2.5. They reveal that a one-unit increase in border controls decreases the individual likelihood of leaving the United States by 10 to 20 percentage points, according to the estimator used. Since each year 55% of the population of U.S. resident illegal migrants

Table 2.7: Effect of a marginal increase in current enforcement on the likelihood of returning from a migration

Model	Static	Lag and lead of bp
$\frac{\partial P(r_{it}=1\mid r_{it-1}=0)}{\partial bp_t}$		
WG AR(1)	-.106***	-.515***
	(.031)	(.155)
Clogit	-.199***	-.136*
	(.001)	(.081)

Block-bootstrap standard errors for logit estimates. *, **, ***= significant at 90%, 95%, 99% confidence level.

leave the country to return to Mexico, such one-unit enforcement increase would decrease average returns by 18% or 36%. The associated elasticity (-0.4 to -0.8) is negative, but larger than -1, implying that a marginal increase in linewatch hours is associated with a less than proportional reduction in returns from illegal trips.

The values of the parameters of interest for the transitions from potential to current migrant status and *vice versa* cannot be directly compared, because they refer to two different populations (the set of all potential illegal migrants, in the first case, and the sub-set of actual migrants, in the second). In the present section, the results from the econometric estimations will be discussed separately. In the next section, estimates of the size of the underlying populations will be obtained, permitting the comparison of border enforcement's effect on migrant inflow and outflow.

2.6.3 Survey design and sensitivity analysis

The sample used for the estimation is a panel built from retrospective information provided by Mexican-based household heads interviewed between

1987 and 1998.[28] The first available year is the one for which I have valid instrumental variables, i.e. 1972 (the year in which DEA was created). Hence, the 1972-1990 migration history of individuals interviewed in 1990 is included in the sample, as well as the 1972-1991 information provided by 1991 interviewees and so on.

Information on household heads' migration history is collected interviewing close relatives, when the former person is absent. However, no information can be collected when the whole family is absent, resulting in migrants being under-represented in the sample. If unobserved individuals differ from observed ones in a way that is correlated with border enforcement, the estimated parameters of interest are biased.

One can think of unobserved migrants as being composed of to groups: those who have "just left" and those who have "not yet returned". The likelihood of both migrating illegally and returning from an undocumented trip is some negative function of border enforcement. Hence, for each level of border patrol, the individual who migrates ("just left") has a higher ability than the observed one who stays, while the (observed) individual who returns from an illegal migration is more able than the one who is still abroad ("not yet returned"). The net effect depends on how comparatively sensitive the probability of migrating and returning home are to changes in border enforcement and on the sizes of both prospective and current migrants in each time period, which are not known *a priori.*

I now proceed to the comparison of observed apprehension rates of migrants from the Mexican and the U.S.-based sample (the group excluded from the analysis). As already mentioned, each year a small-group of U.S.-based migrants are tracked and interviewed, following the indication of the Mexican interviewees. This group is discarded from the sample used in the empirical analysis. Figures 2.6 and 2.7 show trends in average annual apprehension rates and proportion apprehended at least once for illegal migrants

[28] Some individuals are interviewed in 1982. The 1978-1982 valid cases from that panel are included as well.

interviewed in Mexico and in the U.S.A., while Table 2.8 provides respective means and standard deviations.

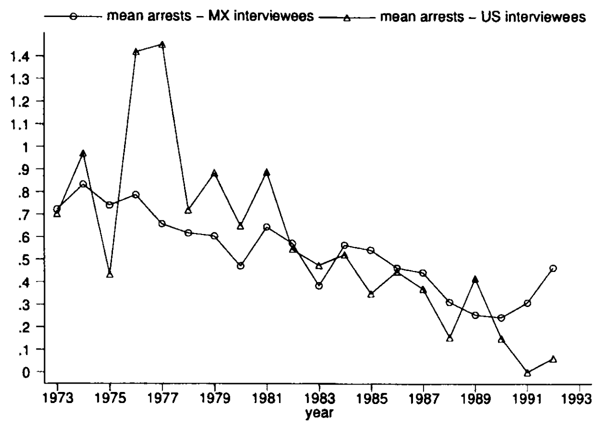

Figure 2.6: average apprehensions for Mexico and US-based interviewees

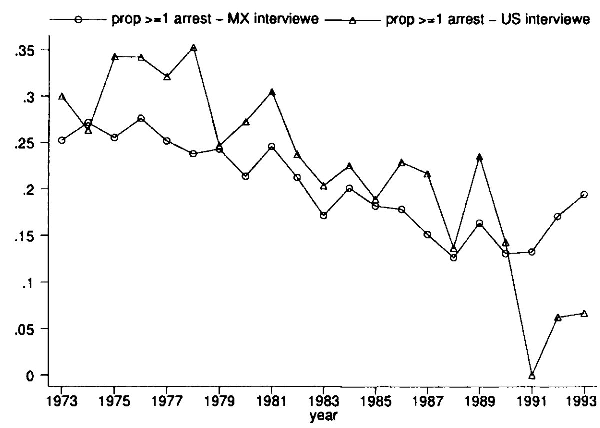

Figure 2.7: proportion apprehended at least once - Mexico and US-based interviewees

Although the annual size of U.S.-based interviewees is quite small, visual inspection of the data does not reveal apparent differences between the two series, with few exceptions. Differences in initial and final years are more marked probably due to the fewer available observations for the U.S.-based

Table 2.8: mean apprehensions and proportion arrested at least once for Mexican- and U.S.-based interviewees, 1972-1993

		Apprehensions		Proportion arrested at least once	
Interviewees:	Observations	Mean	Std. Dev.	Mean	Std. Dev.
Mexican-based	5207	0.532	1.483	0.233	0.423
U.S.-based	982	0.592	1.578	0.256	0.436

group. Table 2.8 confirms that no major difference between the two groups can be detected. Indeed, there is no statistically significant difference between both average apprehension rates and proportion apprehended at least once in the three groups. This result suggests that estimates of the effect of border enforcement on the likelihood of migrating illegally returning from an undocumented migration may not be biased by the absence of some illegal migrants from the MMP71 sample. The result has to be interpreted with caution, given the non-randomness of the U.S.-based sample.

2.7 Interpretation and policy implications

In order to measure the impact of border enforcement on net illegal migration flow, an estimate of the amount of annual migration must be computed. The latter can be obtained by comparing the MMP71 and total border apprehensions, provided the following assumptions hold. First, the observed average apprehension rate is representative of the rate for all illegal Mexican migrants. Second, sample apprehension non-responses are random. Third, non-Mexican migration is constant (proportionally to Mexican ones) throughout time and negligible (in 1996, 97 percent of total apprehended border crossers were Mexican nationals, and this figure includes also apprehensions along the Northern U.S. border), or average border-crossing ability is identical among Mexican and non-Mexican migrants.

Then, knowledge of sample average annual apprehension rate $\overline{a} = a/m$, where a is total sample apprehensions and m is sample undocumented migrations, and of aggregate yearly apprehensions A permits to estimate the annual illegal migration volume:

$$M = \frac{A}{\overline{a}}$$

The corresponding estimated volume of illegal migration ranges between 234,000[29] and 2,570,000 individuals in the years between 1972 and 1993, with an average of 1,265,000. The estimated volume of illegal Mexican migration is consistently lower than the one computed by Massey and Singer (1995) using early waves of MMP data: it ranges between 35 and 77% of their estimates. Early MMP households were sampled from high-migration communities. Villages were migration was a relatively new phenomenon were added later, and are included in the sample (unlike Massey and Singer's) because I use retrospective information. I suspect that this is the reason behind their higher migration estimates. Still, the trends of the two series are not dissimilar, although migration growth in the first half of the 1980s is more marked using Massey's and Singer's figures.

The means for an alternative robustness check of the estimated variable is provided by Espenshade (1995b), who notes how there is a roughly stable proportion between undocumented migrations and total border apprehensions between 1977 and 1988, with a ratio of 2.2. The resulting migration flow computed using aggregate apprehensions would range between 1,075,000 in 1978 and 1,303,000 in 1988, with a mean of 1,323,000. Mean undocumented migrations for the same period using the current estimates amounts to 95% of the Espenshade-based figure. The similarity between the two series is striking, and the difference might be entirely imputed to non-Mexican illegal migration. This comparison is consistent with the hypothesis of small or negligible sample selection in the illegal migrant sample used in the current analysis. Moreover, it indirectly suggest that the MMP71 data may be representative of Mexican illegal migrants at large.

[29] All figures have been rounded.

As an additional robustness check, average annual illegal migration inflow is computed omitting the 1987, 1992 and post-1993 survey years (corresponding to important policy changes), on the grounds that the interviewees in those years may be non-randomly selected. The two series are nearly identical up to 1987, and differ between 1988 and 1991, when, however, there are few available observations for the smaller sample. This is a further confirmation of the sample reliability.

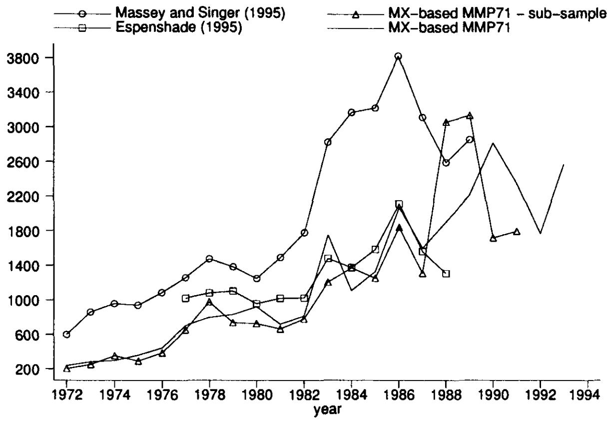

Figure 2.8: comparison of various estimates of illegal migration inflow (in thousands)

Aggregate returns can be calculated by computing the annual departure rate of the sample of current illegal migrants in the United States and comparing it to the estimated illegal migration volume. The population of U.S.-resident undocumented Mexican migrants is extremely mobile. In fact, annual illegal departures for the observed years amount to 95% of undocumented annual entries for the MMP71 sample, corresponding to 1,202,000 mean yearly departures. INS estimates can be used to check the consistency of the estimated variable. The INS calculates that the population of U.S.-based Mexican illegals between 1988 and 1992 grew by 154,000 units per year. Departures from the MMP71 sample constitute 95% of entries for the corresponding years, while average undocumented migration for the

Table 2.9: estimated impact of average annual border enforcement growth on illegal Mexican migration flow, 1972-1993

		Change in inflow		Change in outflow		Net change
Impact of:		(1)[a]	(2)[b]	(3)[a]	(4)[b]	(5)=(1)-(3)
1 extra policeman	LPM	-342	(-205; -477)	-238	(-95; -384)	-104
	Probit/CLogit	-605	(-471; -738)	-457	(-454;-461)	-148

Note: column a's estimates are computed using the estimated parameters of interest from the static models. Column b's estimates are obtained from the 2 standard deviation interval around the respective coefficients. Column (5) is the difference between columns (1) and (3). The first row reports estimates from the linear probability model, the second one from the probit/conditional logit estimates.

1988-1992 interval is 2,208,000. Hence, its 5% is 110,500, not too dissimilar from the corresponding INS estimate. Moreover, sample average net inflow amounts to 87,000 individuals, strikingly similar to the 90,000 average annual inflow computed considering the estimated size of undocumented U.S. residents of Mexican origin (1.1 millions in 1980 and 2 millions in 1990).

I estimate the magnitude of the marginal affect using the enforcement elasticities from the static specifications and considering the impact of hiring an additional patrolling agent. These results are presented in Table 2.9.

Consider the case of an employee working 40 hours per week, 50 weeks per year. This translates into 2,000 additional yearly linewatch hours. Such increase implies a drop in inflow equal to 342 to 605 people and a reduction in outflow of 238 to 457 individuals.

It is worth stressing the multiple conclusions from the above exercise. First, gross undocumented migration greatly overestimates the size of the net one, as migrants are highly mobile and a large proportion of migrations are temporary. It has been estimated that, on average, nearly 1,265,000

Mexicans manage to cross the U.S. border each year between 1972 and 1993. At the same time, migrant outflow is nearly as large, averaging 95% of annual inflow in the MMP71 sample. This point has long been recognized by the existing literature (see, for instance Massey and Singer (1995)).

Second, border patrol's deterrent effect on prospective undocumented migration is sizeable. Given the estimated values, and assuming linearity for a sufficiently large interval around the mean marginal effect, the average annual increase in border enforcement of 77,500 linewatch hours discourages 13,000 to 23,000 individuals per year from crossing the border illegally between 1972 and 1993.

Third, the size of the deterrent effect is substantially reduced (decreased by 70 to 75%) by the induced lower mobility of current undocumented migrants: some illegal resident aliens are dissuaded by the enforcement increase from returning home, since the tougher controls rise current and future migration costs, and lengthen their stay in the host country.

Nevertheless, the influence of border controls is still an overall decrease in the annual flow of undocumented Mexican migration, although quite small in size. The net effect of the average annual increase in linewatch hours is a reduction in net inflow of 4,000 to 5,600 people. The effect of hiring an extra patrolling agent is a decrease in migrant net flow of 104 to 148 units. This translates into high marginal costs to reduce the net flow of illegal migration. Assuming a gross cost of $50,000 for the employment of an extra patrolling agent, the associated net flow reduction of 104 to 148 migrants implies a unit cost of 340 to 480 dollars.[30] The cost is even higher if one considers that on average only 59 percent of work time is spent patrolling the border (GAO, 1996)

The outcome of this analysis raises interesting policy issues. Tight border enforcement seems to involuntary create a more and more permanent population of illegal resident aliens. This might hinder the attempt of erad-

[30]This computation ignores the benefit of the patrolling agent in apprehending drug smugglers.

icating illegal migration even more than having a larger number of highly mobile undocumented workers. The establishment of U.S.-based illegal migrant enclaves may decrease migration costs for prospective migrants (short-term migrants may lack both possibility and interest in helping others out). Moreover, participation to welfare programs might be a positive function of time spent in the United States. This additional indirect cost should be considered in any welfare analysis of the impact of border enforcement policies.

2.8 Conclusions

This paper contributes to the existing literature on Mexican illegal migration to the United States and its border enforcement. It recognizes the limitation of studying the impact of the policy on the inflow of undocumented migrants only by proving that border controls affect the behaviour of both prospective and current illegal migrants, resulting in an ambiguous net effect. Predictions on the effects that increases in enforcement have on both inflow and outflow of illegal Mexican migration are derived by modelling the repeated decision of undertaking illegal trips to the United States. It is shown how rising enforcement intensity has a deterrent effect for prospective illegal migrants because it makes the trip more costly, yet these same higher costs (both current and future) lengthen the U.S. permanence of current undocumented workers.

I present a model were agents may prefer shorter, multiple migrations to a single longer one. The model adds to the existing literature in two ways. First, it endogenises migration duration in a repeat migration setting. Second, it generates optima with multiple spells of migration without introducing shocks or imperfect information. The latter ones are the normal assumptions made by the related theoretical literature to justify the fact that individuals choose to bear the sunk cost of migration more than once. I provide alternative explanations for the observed multiple migrations based on changes in labour demand and on the hypothesis that absence from home bears a positive cost that grows over time. I concentrate on the latter, explaining this cost both in terms of the positive probability of losing claims on current or future asset ownership, and in terms of migrants feeling nostalgic or homesick. Both explanations are consistent with the large observed level of remittances and by the fact that the probability of returning form a migration is a positive function of the number of offspring and of the hectares of land owned by the household.

The necessity of looking at both inflow and outflow of migrants has not been sufficiently acknowledged by the literature, partly because of the scarce availability of direct information on undocumented migrants. Such difficulty is overcome here by merging data on individual repeated illegal migrations with unpublished records of linewatch hours along the Mexico-U.S. border.

The impact of border controls on the behaviour of both prospective and current illegal migrants is then estimated modelling the likelihood of undertaking an undocumented migration and of returning from an illegal U.S. trip. I provide some evidence to support the robustness of the obtained estimates, despite the failure to observe current illegal migrants in the MMP71 sample.

I calculate the volume of illegal Mexican immigration in order to compute the marginal effect of border policing on the net inflow of migrants, comparing sample average apprehension rates with INS data on annual aggregate arrests. The produced estimates are consistent with those obtained applying Espenshade's methodology (1995b). The resulting figures suggest that border controls reduce the total flow of undocumented U.S. foreign workers. The annual decrease of illegal immigrations caused by a marginal increase in border policing is roughly twice as large as the reduction in outflow due to current U.S. illegal residents lengthening their permanence abroad. I estimate that the observed annual growth of 77,500 linewatch hours is associated with a decrease in the illegal migrant net inflow of 4000 to 5600 units, and that an extra patrolling agent results into a net inflow lower by 104 to 148 individuals.

I present some back-of-the-envelope measure of the cost of reducing the net flow of illegal migration. I estimate that an extra 340 to 480 dollars were needed to reduce net flow by one unit between 1972 and 1993.

These findings have interesting policy implications. First, they show that if the intensity of enforcement is chosen without considering its impact on the U.S. permanence of current illegal migrants, it results in too high levels of border control. This may contribute to explain the disproportionate resource allocation to border versus interior enforcement. Second, the findings suggest that tight border policing may contribute to the creation of a more permanent unauthorized U.S. resident population. More research is needed to understand border control's comparative cost effectiveness.

Chapter 3

Aid and Migration: an analysis of the impact of Progresa on the timing and size of labour migration

3.1 Introduction

Migration from developing to industrialized countries has been increasing in the last few decades, both in Europe and in America. Efforts are made in developed countries to limit the migrant flow, composed mainly of unskilled labour and in most cases entering the destination country through illegal channels. Individuals migrate when the associated benefits exceed both its direct and opportunity costs. Migration policies in developed countries aim at discouraging unskilled and illegal migration primarily by decreasing its benefits (from employers' sanctions for hiring illegal labour to reduced access to welfare programs for broad categories of illegal and legal aliens) or by increasing its direct cost (through measures ranging from border enforcement to lengthy or costly visa application procedures).

One possible alternative to discourage migration is to increase its opportunity cost by making the prospective migrant better off at home. One mean to achieve such objective is by channelling resources through aid programs. The relationship between aid and migration is complex, though. On the one hand, transfers are expected to improve economic conditions at home, reducing the economic disparity with the destination locations, hence lowering the incentives to leave. On the other hand, if the observed migration level is inferior to the desired one because of financial constraints, subsidies may be used to fund new trips.[1] Aid programs will affect the likelihood of migrating by changing both income differentials and the possibility of financing trips.

In the last few years a specific type of aid policy has been adopted by a number of Latin American countries. This program, called Progresa in Mexico, is aimed at improving education, health and nutrition of poor rural households. The objective of the current analysis is to understand the relationship between this program and migration, both intra- and international. The focus of the analysis is on Mexican migration and it is dictated by data availability. However, the conclusions reached may be easily extended to other sets of countries.

Progresa is an ongoing program that targets poor Mexican rural households (and has been recently extended to urban locations under the name of Oportunidades). Among the various components of the program, there are a (smaller) unconditional nutrition support grant, and some (cumulatively larger) schooling subsidies, conditional upon attendance of the last four grades of primary school and first three grades of secondary school. All grants are paid bimonthly, and the schooling subsidy is received upon proof of attendance of at least 85 percent of classes. The size of the transfer grows with the school grade, and is estimated to correspond to two thirds of the wage earned by a teenager in full time employment (Schultz, 2004).

I argue that, although similar in type, primary and secondary school sub-

[1]The relationship between aid and migration for households facing liquidity constraints and having a home bias has been explored, among others, by Faini and Venturini (1993, 1994, 2001). Examples of the use of preference for home consumption to justify return migration may be found in Djajic and Milbourne (1988) and Dustmann ().

sidies have different implications in terms of their impact on the recipients' time allocation. Indeed, the subsidy to primary school attendance is roughly equivalent to an unconditional transfer, since primary school enrolment is very high. The secondary school grant conditionality constraint is instead binding for the majority of potential recipients, as secondary school enrolment is much lower. Some eligible schoolchildren are potential migrants. About 20 percent of the observed international migrations occurring in the sample are undertaken by individuals aged 13 to 19.

I study the impact that Progresa has on labour migration only, defined as the act of leaving one's hometown to seek employment elsewhere. Progresa enters the decision process of prospective migrants in beneficiary households by changing their opportunity costs, their financial resources and their expected wages.

I group the three monetary components of Progresa into unconditional and conditional transfers. The former refers to the nutrition support and the primary school grant; the latter is the subsidy to secondary education. I show that the two types of grants have different impacts on migration. The unconditional transfer may reduce migration, if individuals have a preference for consumption at home or, more generally, if migration (intangible) costs depend positively on household income. This effect regards the least poor households. At the same time, the unconditional grant may result in higher migration levels because it relaxes financial constraints for the poorest households. The latter fact may have sizeable effects, given that the imperfection of capital markets is one acute problem faced by indigent families in developing countries.

The conditional transfer may reduce migration in the short term by requiring recipients to stay in the home country: prospective migrants may be deterred from migrating, provided that the size of the grant is high enough. Whether beneficiaries revert to migrating at the end of the program depends on skilled wages at home and in the destination locations. In addition, the secondary school subsidy may provide migration incentives to individuals who would have not left home before the program implementation. The short-run impact of the two types of transfers on migration are estimated separating the pure income effect from the conditionality one. I distinguish the effect on domestic and international migration.

An interesting feature of the program is that 186 of the 506 sampled villages are randomized out. Eligible residents of these villages are not administered the program until 2000. Eligibility is determined by poverty status, which is defined on the basis of household-specific information collected in the pre-program September 1997 census of Progresa localities. All residents of both control and treatment villages are then interviewed at regular intervals. Detailed data on migration are collected on an annual basis. This provides us with information on households in both control and treatment villages observed both before and during the implementation of the program. I exploit the exogenous variation induced by the randomization to obtain a valid counterfactual for treated households, in order to assess the impact of the various program components on migration.

The empirical analysis confirms that unconditional cash transfers are associated with increased migration, while secondary school grants reduce short-term migration.

The structure of the paper is the following. Section 2 uncovers the relationship between migration and aid, in the form of both unconditional and conditional transfers, by sketching a model of migration and schooling choice. Section 3 describes the data used in the empirical analysis. Section 4 describes the specification used and discusses the identification of the estimable parameters. Section 5 presents the results. Section 6 comments on some policy implications and concludes.

3.2 Aid and migration: theoretical considerations

Mexican migration, especially international one, tends to be temporary in nature, rather than permanent. Household resources are pooled to finance the migration of one or more of its members, normally young males, who leave their family in the community of origin and spend time away, remitting money at regular intervals, or bringing back their savings with them. Remittances are often a sizeable proportion of migrant households' income.

Virtually all children attend primary school before the program implementation, as shown in the next chapter. Hence, since most kids would have attended school even without the transfer, the eligibility constraint is not binding and the subsidy has a pure income effect for the near totality of households.

This is not the case for secondary schoolchildren. While re-enrolment rates are quite high for kids with some secondary schooling, the likelihood of beginning high school is low: before the program, less than 40 percent of children aged up to 16 with complete primary school enrol to high school (Attanasio *et al.*, 2001). Three possible factors may explain the low transition rate from primary to secondary education, which are both free in Mexico. One is the higher distance from school, not often available in the village of residence (unlike primary schools, present in most localities). The second one is the higher opportunity cost, as forgone earnings are likely to be higher for teenagers than for younger children. This is especially true given that teenage offsprings are potential migrants. In November 1998, one third of all migrants from poor households in the control group were up to 19 at migration. 78 percent have some primary school education. 50 percent of migrants have completed primary school when they migrate. A further 3 percent has some secondary education, while 16 percent has completed junior high school.

Progresa's schooling grant is conditional upon school enrolment and attendance of at least 85% of classes. Thus, the constraint for transfer eligibility is binding for the majority of households with secondary schoolchildren. Given the differences in pre-program enrolment rates, I consider both income support and primary school grants as unconditional trasnfers, and the secondary school subsidy as a conditional one.

The effect of a transfer to poor rural households on migration is going to differ according to its eligibility requirements. The effect of an unconditional transfer for these families may be a decline in migration, provided that agents sufficiently dislike moving abroad (which, in turns, may be a function of their wealth). However, the transfer may relax the financial constraint faced by poor households: if families could not reach the desired migration level before the program because of borrowing constraints were coupled with the impossibility to save, they may use the cash subsidies to fund additional trips. Overall, then, the net effect of the unconditional transfer depends on wealth and on access to the credit market, as well as on the "standard" variables from any migration model, and cannot be predicted *a priori.*

In the case of the conditional school subsidy, if household consumption is sufficiently far from subsistence level, and if the returns to secondary education are high enough, children may move from employment (and migration) to schooling. Grant size for grades 7 to 9 amounts only to approximately two thirds of the wage a child of the corresponding age might earn if working in the village of residence[2] (Schultz, 2004), however secondary education is expected to provide access to better-paid jobs. The short-term migration reduction might be offset by an increase in the medium run, after secondary school is completed. Once the education cycle is completed, the individual will move to the location that pays highest relative wages, net of moving costs.

3.2.1 Effect of an unconditional transfer

The above considerations can be illustrated with the help of a simple model. The effect of an unconditional transfer can be captured in a static framework. Assume there is a continuum of poor households whose utility (u) depends positively on consumption. Households' pooled resources are labeled $Y \geq 0$ (income) and are continuously distributed along their support, with some

[2]The size of the transfer rules out the possibility that a child is enrolled in secondary school purely to finance the migration of a family member, if the likelihood of finding a job is very high.

density function $F(Y)$. Suppose as a starting point that the only source of heterogeneity across households is $Y \geq 0$.[3]. Household members include one child who attends one of the subdisied primary school grades and another one who may be eligible for the secondary school grant. There are two locations, "home" and "away". While the younger child attends school and does not work, the elder may choose between working at home ($m = 0$), earning a wage w^h, and migrating ($m = 1$), where a wage $w^a > w^h$ is earned. I rule out any form of intra-household bargaining for the allocation of resources or of strategic behaviour by assuming that all members care equally for each other.

Migration has a positive cost. This cost is composed by a monetary element K (such as travel expenses), which is exogenous to the migrant, and by an intangible one, which is a positive function of potential home consumption. Assume for simplicity that intangible migration costs equal $(Y + w^h)/\alpha$, with $\alpha > 1$. The wealthier the household, the larger is the intangible cost of not being together. This can be thought of as individuals suffering from being (or consuming) apart (as supposed in the previous chapter). $1/\alpha$ measures the comparative dislike for "away" consumption, or for not consuming together at home. The smaller the fraction, hence the larger α, the smaller is the disutility from consuming away from home. When $\alpha \rightarrow \infty$, household members are indifferent between the location where consumption occurs, and care only about its maximization.[4]

I represent financial constraints by ruling out saving and borrowing. All income earned is consumed in the current period. I further assume that the migration cost is borne up-front, at the beginning of the period. Hence, only households with sufficiently high income ($Y \geq K$) are able to finance the migration, since there cannot be negative consumption. I also assume for simplicity that utility is a linear function of the parameters of the model.

[3]This assumption is not required in the empirical analysis

[4]The latter is a broad version of the standard opportunity cost argument, including all types of goods that cannot be enjoyed while abroad.

The maximization problem is

$$\max_{m=\{0,1\}} u = Y + (1-m)w^h + m(w^a - K - (Y + w^h)/\alpha)$$

All households may potentially benefit from the migration, but only those with income such that $Y \geq K$ and $Y < (w^a - w^h - K)\alpha -\ w^h$ will undertake it. Very poor families (with income $Y < K$) are unable to finance the trip, while the least poor ones ($Y > (w^a - w^h - K)\alpha -\ w^h$) have too high an intangible cost. They are already sufficiently wealthy that the higher income associated with the migration is offset by the cost of being apart.

An unconditional transfer (T) in the form of a nutrition support or a primary school subsidy has a twofold effect. It increases household income, relaxing the financial constraint of some of the families for whom migration was precluded in the baseline case. Hence, some households who were previously credit constrained will be able to pay for the migration. These are families with income $Y \in [K - T; K)$. At the same time, the transfer increases migration costs, $K + \frac{(Y+w^h)}{\alpha} < K + \frac{(Y+w^h)+T}{\alpha}$. Households with incomes close to the upper threshold will stop migrating. This holds for families with income $Y \in [(w^a - w^h - K)\alpha - w^h - T; (w^a - w^h - K)\alpha - w^h)$.

The overall effect is ambiguous and depends on $F(Y)$. However, the more stringent the financial constraint and the poorer the households, the higher the likelihood that there will be a net increase in migration. [5] [6]

3.2.2 Effect of a conditional transfer

I now proceed to illustrate the effect of a secondary school subsidy. The nature of this problem is dynamic, because current education is associated with higher future earnings. I capture the dynamic effects by adding a second period to the model and by modelling explicitly the schooling choice

[5]In the absence of intangible migration costs linked positively to income, the effect of the unconditional transfer is an unambiguous increase in migration.

[6]Modelling temporary migration by letting individuals leave for a fraction of the period (i.e. $m \in [0,1]$) yields the same ambiguous impact of unconditional transfer on migration flow, while it reduces migration duration.

of the adolescent son. This choice is discrete ($s = 0, 1$), hence school can be attended for the whole period only and is not compatible with migration. School attendance has a positive price, p, which represents both its direct (e.g. travel expenses) and its opportunity cost (forgone earnings). Given the borrowing constraints, only households whose income exceeds education costs ($Y > p$) are able to finance education. Each pupil is exogenously endowed with the ability to transform schooling into particular types of skills, which are rewarded differently at home and away. Hence, each student in the second period has pairs of home and away-specific human capital, $k = (k^h, k^a)'$, which will determine wages in both locations ($w^h(k)$ and $w^a(k)$). Schooling is associated with skilled wages, $w^j(k) > w^j$, for $j = h, a$ and for all k. Wages are identical in the two periods for individuals with no secondary education. Education changes the incentives to migrate by increasing next period home and away wages and by raising the intangible cost of migration (through the higher home wage). The intensity of these effects is a function of individual-specific returns to education, and the decisions will vary depending on households' income levels.

The maximization problem becomes

$$\begin{aligned}
& \max_{m=\{0,1\},s=\{0,1\}} u_1 + u_2 \\
u_1 &= Y + (1 - m_1)(w^h - ps) + m_1(w^a - K - (Y + w^h)/\alpha) \\
u_2 &= Y + (1 - m_2)w^h(k) + m_2[w^a(k) - (K - (Y + w^h(k_{1\{s=1\}}))/\alpha)_{1\{m_1=0\}}]
\end{aligned}$$

There are six potential outcomes, depending on whether one migrates or not in either period and on the schooling choice in period one. The associated utility is represented by the triplets (m_1, m_2, s), where each parameter can take two different values (0 or 1). Hence, a household where the teenage son goes to school in the first period and migrates in the second one has a utility of $(0, 1, 1)$. Different outcomes are chosen according to family income and returns to education at "home" and "away". These are summarized in Table 3.1, and discussed below.

Table 3.1: Inter-temporal schooling and migration decisions

	$Y < p$	$Y \in (p, K)$	$Y \in (K, \bar{Y})$	$Y > \bar{Y}$
Low k	$(0,0,0)$	$(0,0,0)$	$(1,1,0)$	$(0,0,0)$
High enough k^h	$(0,0,0)$	$(0,0,1)$	$(0,0,1)$	$(0,0,1)$
High enough k^a	$(0,0,0)$	$(0,0,0)$	$(0,1,1)$	$(0,0,0)$

The poorest families with $Y < p$ will neither have their teenage child migrate nor go to school. Those with income $Y \in (p, K)$ will choose whether to continue education depending on the level of skilled home wages. Consider a deterministic world where individuals know their human capital production function. Agents will go to school if the following disequality holds: $w^j(k) - w^h > p$ for $j = h, a$ if $Y \geq K$ and for $j = h$ if $Y < K$. It follows that for households with income $Y \in (p, K)$, $(0,0,1) > (0,0,0)$ for sufficiently high k^h.

Since $w^a > w^h$, agents who choose not to purchase education will migrate in both periods, provided that they are not financially constrained. This occurs because once the migration cost has been borne, consumption is maximized by staying away for as long as possible. This implies that $(1,1,0) > (1,0,0)$ and $(1,1,0) > (0,1,0)$. Hence, migrants who are not financially constrained will choose one of the three outcomes: $(1,1,0)$, $(0,1,1)$ or $(0,0,1)$ according to their location-specific human capital. For instance, outcome $(1,1,0)$ will be chosen by households with pairs of human capital k and income Y such that

$$2w^a - K - (Y + w^h)/\alpha > w^h(k) - w^h - p \tag{3.1}$$

i.e. $(1,1,0) > (0,0,1)$, and

$$2w^a > w^a(k) + w^h(\alpha + 1)/\alpha - w^h(k)/\alpha - p \tag{3.2}$$

i.e. $(1,1,0) > (0,1,1)$.

Three classes of individuals do not migrate in either period. Those with very low income, $Y < K$; those with high enough k^h to make "home" skilled wages preferable to migrating in one or two periods; the wealthiest households. This latter group is composed by families with income $Y > \bar{Y}$, where

$$\bar{Y} = [w^a(k) - p - w^h - K]\alpha - w^h(k)/\alpha \tag{3.3}$$

for individuals with high k^a (whose alternative is to go to school in order to access the "away" skilled wage $w^a(k)$) and

$$\bar{Y} = [2(w^a - w^h) - K]\alpha - w^h/\alpha \tag{3.4}$$

for those with low k (whose second best is to migrate in both periods).

The education subsidy reduces the price of education by an amount Δp. The general effect of the subsidy of course is to increase attendance of secondary education. Regarding its effect on migration, the schooling subsidy unambiguously reduces short-term migration. This happens because the lower cost of education induces some individuals to move from a two-period migration to education. Thus, the subsidy is associated with a reduction in the number of $(1, 1, 0)$ choices in favour of $(0, 0, 1)$ and $(0, 1, 1)$ outcomes.

The subsidy's medium-term effect on migration is instead uncertain. In fact, individuals with sufficiently high "home" return to education will not leave in the second period, i.e. they will switch from $(1, 1, 0)$ to $(0, 0, 1)$, as shown by (3.1). Agents with a large enough "away" skilled wage will instead go to school in the first period in order to migrate in the following one and to access "away" skilled wages, moving from $(1, 1, 0)$ to $(0, 1, 1)$, as illustrated in (3.2). In addition, the lower price of education will induce children in some of the least poor families (the ones with high "away" returns to education) to attend secondary school in order to migrate in the second period, moving from $(0, 0, 0)$ to $(0, 1, 1)$, as noticeable from (3.3).

Although the net effect of the secondary school grant on medium-term migration is not clear, it is possible to make some general considerations to understand its potential overall effect. Given that the program targets poor rural households, it is expected that "home" returns to higher education are not as high as "away" ones. It is probable that individuals who switched from migration to education in the first period will then leave the village in the second one. However, multiple migration destination are possible, depending on comparative net benefits from a domestic migration to urban areas and an international one. The comparative advantage of an international versus a domestic migration is lower for educated than uneducated migrants, if individuals can only enter the U.S. illegally, since it is likely that wages in the secondary segment of the U.S. labour market are less sensitive to changes in migrants' education level. Considering in addition the higher migration cost, it is not certain whether migrants with secondary schooling will migrate abroad or remain in Mexico.

In short, the effect of Progresa on migration cannot be inferred with certainty. This is due various facts. First, the program's different components may provide opposite incentives to migrate. Second, the same type of subsidy may affect household's migration level differently according to their degree of poverty, of credit constraints and their children's comparative skilled wages at "home" and "away". Third, the effect of the conditional component of the program may vary in the short and in the medium run.

Modelling the effect of Progresa on migration has highlighted the following points. The unconditional grant links current migration to poverty by an inverse u-shaped function: it increases emigration of some poor households by relaxing their credit constraints, while reducing the incentives to leave of the least indigent recipients. Given the poverty level of program recipient, though, one expects that the loosening of financial constraint associated with the program recipience dominates the higher intangible migration costs, resulting in a net migration increase. Contemporaneous migration appears to be a negative function of the conditional grant, instead, as both the high school subsidy and the future access to skilled wages provide incentives to stay at home. Hence, I expect the conditional and the unconditional program components to have opposite effects on the likelihood of migrating.

The volume of future migration depends positively on "away" skilled wages and negatively on the "home" one. If the latter is sufficiently high, some individuals will refrain from migrating also in the second period. However, given the comparative scarcity in skilled labour demand in rural areas, one may expect migration to increase in the second period. Whether individuals will migrate to Mexican urban areas, or will prefer to go to the United States depends on the comparative net benefits, which are functions of wages and costs, which in turn depend on the type of international migration (legal or undocumented).[7]

The estimation of the observed net impact and the effect of the program's various components is the subject of the next sections.

3.3 The Progresa data

Progresa targets Mexican poor rural households and provides grants to improve education, health, consumption and the role of women in the household. Its main monetary component, apart from a smaller nutritional subsidy, is in the form of a schooling subsidy to children attending the last third grades of primary school and the three grades of secondary school. Transfers are made to women only (normally the spouse of the household head), and are conditional upon regular visits to health centres, to "*platicas*" where women are taught about health and nutrition issues, and to a school attendance rate of at least 85% of term time. An interesting feature of the program is that, in order to permit the evaluation of its impact, 186 of the 506 villages sampled for evaluation purposes are randomized out. Poor residents of these villages are not administered the program until 2000. Households are classified into poor and non-poor according to the information collected in the pre-program September 1997 census of Progresa localities.

[7]Progresa is likely to influence migration through two further channels: changes in income distribution and risk. Transfers to poor households reduce both their relative deprivation (i.e. their rank in the community income distribution) and income variation. Since migration is a mean to smooth risk and improve one's relative, as well as absolute income, then Progresa will result in an unambiguous migration reduction through these channels.

All residents of both control and treatment villages are then interviewed at biannual intervals. Detailed data are collected on health, consumption, income and employment, education and migration at least in one of the two annual surveys.[8]

One shortcoming associated with the implementation of Progresa is the presence of a group of eligible households who were "forgotten". This means that a proportion of households randomized in the program did not receive the due benefits initially.[9] The likely explanation for this fact is attributed to administrative delays and mis-classifications. These were likely caused by the dual round of selection of eligible households. In fact, a group of households initially classified as non-poor were later included in the beneficiary group. This process, known as *densificacion*, took process at the beginning of 1998 and increased the proportion of eligible households in treatment villages from 52 to 78 percent.

The absence of an observable counterfactual for "forgotten" eligible households complicates the estimation of the program impact on treated subjects. It is studied in the next chapter, where I show a set of conditions under which it is possible to estimate the average program effects for all households in the Progresa sample. However, such approach is not pursued here because of the smaller sample sizes of the relevant groups and the very few migrations occurring within each category. Thus, I restrict my sample to the sole households classified as poor in 1997, who constitute 52 percent of the sample, as already mentioned. Further details on the data are provided both in the next sections and in the following chapter.

I consider labour migration only. Both the 1998 and the 1999 surveys (unlike the 1997 one) record the motivations for leaving one's household of origin. I distinguish between domestic and international migration. I will refer alternatively to international or U.S. migration, since this is the

[8]Information on migration is collected in the September 1997, October/November 1998 and November 1999 waves, i.e. rounds 1, 3 and 5 of the survey.

[9]The difference with the "standard" dropout issue is that here the choice of not being treated is exogenous for "forgotten" households.

most likely destination for international migrants. 95 percent of all trips occur when the individual is aged between 14 and 40. Thus, I consider the sole subset of people within this age interval as potential migrants. Older and younger individuals are discarded from the analysis. As a result, entire households are dropped from the valid sample, which is composed of approximately 27,000 individuals from 10,000 households. About 17,000 individuals (7,000 households) belong to the treatment group, as shown in Table 3.2.

Table 3.2: Sample size of sub-groups

	1997	1998			1999		
	All	B	F	NC	B	F	NC
Treatment							
Individuals	16,877	16,532	27	594	15,590	25	464
Households	7192	6617	6	260	6136	5	194
Control							
Individuals	10,278	10,295			10,029		
Households	4314	4096			3909		

B=beneficiary; F=forgotten; NC=non-complier. 1998 data.

The treatment group in 1998 and 1999 is split according to transfer recipience: B stands for actual beneficiaries, while F and NC indicate forgotten and non-complier subjects, respectively. Only 6 families in 1998 are not administered the subsidies, although entitled to, irrespective of compliance with program requirements. The number of non-compliers is larger, but still quite small when compared to the size of the treatment group. Indeed, the sum of the two groups who do not receive the transfers among 1997 poor amounts to around 3 percent of total potential beneficiaries in the post-program years.

3.4 Econometric analysis

The small size of non-compliers and the absence of pre-program significant differences in average migration rates between control and treatment group simplify the econometric analysis greatly. The major advantage of the randomization is that control and treatment group do not differ in terms of unobservable characteristics. Hence cross-sectional analyses can provide consistent estimates of average program effects. Although group means provide unbiased estimates of the program effect, regression analysis may be performed to increase the estimate precision and to control for relevant variables (such as the presence of shocks, which may not occur randomly in case of natural disasters occurring in specific geographic areas).

3.4.1 Estimable parameters

Two parameters commonly estimated in the program evaluation literature are intention to treat and average treatment on the treated effects ($TTEs$). The former is estimated by comparing all eligible individuals in the treatment and control groups. This parameter can be interpreted as measuring the average program effect for all eligible households irrespective of actual treatment.

$$E[m|T=1] - E[m|T=0] \tag{3.5}$$

$T = \{0, 1\}$ for subjects in control and treatment villages, respectively. m is some measure of migration to be discussed later.

TTE measures the average impact of the program on actual beneficiaries. Since non-compliers are quite scarce in the considered group, amounting to approximately 3 percent of the sample, I expect the difference between treatment availability and use to be negligible.[10]

[10] The results from the empirical analysis do not change if these two groups are omitted.

$$
\begin{aligned}
E[m|T &= 1] - E[m|T=0] \\
&= 0.97\left\{E[m|T=1,B=1] - E[m|T=0,B=1]\right\} \\
&\quad +0.03\left\{E[m|T=1,B=0] - E[m|T=0,B=0]\right\} \\
&\cong E[m|T=1,B=1] - E[m|T=0,B=1]
\end{aligned}
$$

B refers to being a program dropout, or actual beneficiary, $(=1)$ or not $(=0)$. In the remaining sections I will only refer to TTEs, although the parameters that I am actually estimating are the intention to treat ones.

The actual grant received by beneficiaries is not observed. Not all eligible children may end up going to school. Thus, I will use potential grant size and composition when estimating the effects of the program components. Nevertheless, the resulting parameters have a clear policy significance, since they measure the impact of what is under the control of the policy maker, rather than parameters depending on households' acceptance of the treatment.

3.4.2 Program effect: specification and identification

The identification of average $TTEs$ relies on the village randomization and is based upon the claim that eligible individuals do not differ from control ones in terms of unobservable characteristics. Thus, unbiased counterfactuals of the above parameters are given by their sample analogs ($\bar{m}_T$, $T = 0, 1$). The simplest way, and natural starting point, is to compute simple group means and test for their statistical difference.

$$E[m|T=1] - E[m|T=0] = \bar{m}_1 - \bar{m}_0$$

In the case of a dichotomous variable these are the difference in average likelihood of migrating due to Progresa. I adopt this approach to estimate the overall program effect on migration, alternatively pooling migrations and distinguishing between domestic and international ones.

The theoretical model highlights how the conditional and unconditional components of the transfer may impact migration differently, hence how a grant of a given size may affect households in varying ways, depending on beneficiaries' and households' characteristics. A simple way to test these hypotheses is to condition on potential grant size and composition. I create a variable that measures the potential size of the grant to which households would be entitled, were all their children to attend school. Potential grant size varies between 190 and 230 *pesos* and is capped to a maximum of 1250 and 1500 *pesos* (in November 1998 and 1999, respectively). The increases in grant levels are such that the value of the subsidy is constant over time. I also compute the proportion of potential grant associated with eligible males' secondary school attendance. I consider males only because they are more likely to be migrants than females in the same age group.[11] At most 83% of total potential grant may come from male secondary scholarships because of the income support provided to all eligible families.[12]

I group subjects according to the household grant size (g) and composition (p). g takes two values: 0 for low grant sizes, up to its mean, and 1 for larger than average transfers. p has three levels: 0, for households with all unconditional grant; 1, for up to 50% conditional grant, which is the median for households with secondary male schoolchildren (the mean is 51%); 2, for more than 50% conditional grant. I first test whether there are significant differences in the migration propensity of treatment and control households grouping them by grant size and composition. In other words, I estimate the following effects:

$$\begin{aligned} TTE_{gk} &= E[m|T=1, g=k] - E[m|T=0, g=k] \\ TTE_{pj} &= E[m|T=1, p=j] - E[m|T=0, p=j] \\ k &= 0,1 \qquad j=0,1,2 \end{aligned}$$

[11] However, for robustness checks I perform the empirical analysis considering also grant composition in terms of female secondary school subsidy. The latter does not appear to be related to migration as much as the former.

[12] Further details regarding the creation of these variables, the presence of a common support among control and treatment households and the amount of variation within the two variables are provided in the Appendix.

There are five such effects. Define m_i^* the dependent variable of the latent regression model that determines the migration choice according to the following equality: $Pr(m_i = 1) = \Pr(m_i^* > 0)$, where m_i indicates whether the migration is undertaken or not. In the case of discrete g and p:

$$m_i^* = \alpha_0 + \alpha_1 T_i + \alpha_2 g_i + \sum_j \alpha_{3j} p_{ij} + \alpha_4 g_i T_i + \gamma X_i + u_i \qquad j = 1, 2 \quad (3.6)$$

X represents additional control variables. u is some white-noise error following some fully-specified parametric distribution. The subscript i indicates the *i-th* individual, $i = 1, .., N$. When $g = 0$ and $T = 0$ are the omitted groups and the regression is estimated by linear probability model[13], $TTE_{g1} = \alpha_1 + \alpha_4$ and $TTE_{g0} = \alpha_1$. Inverting g and p provides estimates of the TTEs for the composition effects. These parameters permit to compare different groups of households, but they do not permit to understand whether any difference in migration propensity is attributable to size or conditionality effects.

To disentangle the two effects, I then proceed to estimate the impact of each component, conditioning on the other one, i.e. interacting both by the treatment dummy. Thus, I measure the income effect testing whether, conditioning on grant composition, the effect of the grant on migration is larger for larger grant sizes. For instance, suppose that households face severe financial constraints, and use the grants to fund additional migrations. After conditioning on grant composition and other variables that capture the overall migration propensity in the family, it is expected that migrants with more funds will be more likely to migrate than those with little money transferred to. In other words, one expects the following equation to be positive.

$$\begin{aligned} & \{E[m|T = 1, g = 1] - E[m|T = 0, g = 1]\} \qquad (3.7) \\ & - \{E[m|T = 1, g = 0] - E[m|T = 0, g = 0]\} \\ = \ & TTE_{g10} = TTE_{g1} - TTE_{g0} \end{aligned}$$

[13]For nonlinear models, the coefficient has to be multiplied by the value of the cdf at the specific variable levels.

The same applies to the composition effects, hence the labels TTE_{pjl}, with $j, l = 0, 1, 2$ and $j \neq l$. The subscripts j and l refer to the three possible pairwise comparisons of the conditional grant proportions (p=zero, low and high, as explained above). In the case of discrete g and p:

$$m_i^* = \beta_1 T_i + \beta_2 g_i + \sum_j \beta_{3j} p_{ij} + \beta_4 g_i T_i + \sum_j \beta_{5j} p_{ij} T_i + \gamma X_i + u_i \qquad j = 1, 2 \tag{3.8}$$

The coefficient on interacted grant size from (3.8), β_4, provides an unbiased estimates of (3.7). The coefficients β_{5j} capture the impact of program conditionality on migration: $\beta_{51} < 0$ and $\beta_{52} < 0$ support the view that, given grant size, migration is lower among household where the program requirements are binding.

In all regressions, I add a set of conditioning variables that are expected to capture different migration propensities of households with different demographic composition (a detailed explanation follows in the result section). Furthermore, I estimate (3.8) conditioning on continuous g and p, assumed alternatively to be linear and quadratic. In this case, the above coefficients are measuring the change in migration propensity caused by a marginal increase in the money received (in the case of g) or in the proportion of the grant due to male secondary scholarships. Identification in this case hinges on the explicit functional form assumptions.

The magnitude of the grant's income effect depends on its relative size, when compared with average migration costs. Average monthly grant size for families with children is 348 *pesos*[14], and 250 *pesos* for childless households, while the male secondary school subsidy varies between 200 and 250 *pesos* in 1998 and 1999. I expect costs to vary with distance from the chosen locality, and to be highest for international migrations also for an additional reason: the vast majority of international migrations are illegal and tend to hire smugglers to cross the border (77 percent of illegal migrants resorted to hiring a smuggler in the 1980s and 1990s, paying on average 540 1990 dollars for the years 1993-1998). Given these high U.S. migration costs, it is

[14]One dollar was roughly 10 pesos in 1998 and 1999.

unlikely that the cumulative grant in November 1998 is sufficient to finance an international trip.

The program may provide incentives to return from a migration. Looking at average remittances, the direct financial incentives appear to be low for U.S. migrants, but sizeable for domestic ones. International and domestic migrants remit on average 80 and 400 *pesos* per month, respectively, and households with migrants have on average 1.5 members away.[15] No direct information is available on return migrants in the 1998 and 1999 waves. However, data on migrants' stocks are collected. I order to obtain an estimate of the program impact on migration net flows I compare the stocks of labour migrants in the treatment and control groups. Given the absence of pre-program mean difference in migration levels between control and treatment households, any difference in stock can be attributed to differences in net flows:

$$\begin{aligned} S^T_{98} - S^C_{98} &= S^T_{97} - S^C_{97} + NF^T_{98} - NF^C_{98} = NF^T_{98} - NF^C_{98} \\ \text{if} \quad S^T_{97} &= S^C_{97} \end{aligned}$$

where S and NF indicate migrant stock and net flow, the subscripts refer to the relevant year and the superscripts to treatment and control group. When comparing migrant stocks in 1999, I am looking at the difference in net flows in the two years after the pre-program data are collected.

3.4.3 Does the randomization work? Pre-program means

The identification of the parameters of interest is based on cross-sectional variation. Thus, it relies on the validity of the randomization, meaning that there are no unobservable differences in migration patterns among treat-

[15]Average Progresa subsidy size data are from Albarran and Attanasio (2001), information on smuggler hiring and associated cost has been computed from data collected by the Mexican Migration Project. Migrant number and monthly remittances have been computed from the November 1998 Encel survey, using average values for the control group.

ment and control groups.[16] One possible way to insure that time-invariant unobservable individual effects are not driving the result may be to perform some difference-in-difference analysis. However, labour migrations are not identified in the 1997 data. All household members who are away at the time the interview is carried are classified as migrants, including those who left to get married and to go to study. While Progresa is expected to increase labour migration, the other types of migrants may decrease, given the higher incentives to stay in treated villages. If the two effects offset each other, the program may not appear to affect labour migration, even though it might actually do.

Alternatively, one can test for pre-program different migration levels. Given the interest in the effect of the grant for different household composition, one has to control whether there are significant differences in pre-program migration levels for families with different number and type of schoolchildren. Extensive evidence will be provided of the absence of significant differences in the migration rates of various sets of households in the treatment and control groups.

I classify migration according to its location, computing three different variables (total migration, domestic and international one). For each of these variables, I consider: average individual migration, proportion of families with at least one migrant and average household migration (for households with a positive number of migrants).

Average pre-program migration rates along these three dimensions are presented in Table 3.3. No significant differences are detected between treatment and control subjects. As migration motives are not known in 1997, the values below include not only labour migrations, but also trips for educational purposes (very few ones) and to get married.

Although true labour migration rates are inflated, the number of individuals who leave the household in 1997 is extremely low. Fewer than

[16]This assumption is especially crucial in light of the few migrations occurring in the sample.

Table 3.3: Pre-program mean migration levels

		Valid 1997 observations		
		All	US mig	MX mig
Treatment	Individuals	.0174	.0071	.0103
		(.0019)	(.0012)	(.0014)
	Households (% mig>0)	.0129	.0055	.0076
		(.0016)	(.0010)	(.0012)
	Households (mean) [a]	1.505	1.475	1.472
		(.0730)	(.1299)	(.0848)
Control	Individuals	.0149	.0053	.0095
		(.0025)	(.0011)	(.0023)
	Households (% mig>0)	.0118	.0044	.0076
		(.0019)	(.0011)	(.0015)
	Households (mean) [a]	1.588	1.650	1.484
		(.1424)	(.1596)	(.1902)

[a]computes mean migrant number for household with at least one migrant. Standard errors clustered at the village level.

one percent of the sampled individuals are foreign migrants, the proportion rising only slightly for domestic migration. The low migration rate is consistent with the financial constraint hypothesis: some profitable migrations may not have been funded because of the impossibility to borrow.

Table 3.4: Pre-program migration differences for various percentiles

	Grant size		Composition[a]	
p-value of difference	migUS	migMX	migUS	migMX
25-th percentile	0.678	0.499	0.263	0.886
50-th percentile	0.160	0.719	0.263	0.886
75-th percentile	0.291	0.473	0.263	0.886
90-th percentile	0.229	0.813	0.355	0.345

[a]: grant composition measured as proportion of grant due to male secondary school attendance. Standard errors clustered at the village level.

The computation of TTEs for several sub-groups of individuals and the interaction with family demographic characteristics (as measured by proportion of grant due to secondary school grant for males) means that one needs to test whether the pre-program distribution of migration differs among control and treatment villages. Mean equality in this case is not a sufficient condition. I test for difference migration rates for households at the 25-th, 50-th, 75-th and 90-th percentile of both potential grant size and grant composition measures. The results, which are reported in Table 3.4, show that there are no significant differences in pre-program migration rates. I report results at the individual level only, since they convey the same message as the household-level ones.

To further test for the absence of significant differences in pre-program migration levels, I regress (3.6) computing the pre-program equivalent of TTE_{gs} and TTE_{ps} using 1997 data. All parameters but one are not statistically different from zero.

Table 3.5 reveals that households with a high proportion of secondary

Table 3.5: Pre-program migration difference for various household types

	Individual			Household		
	migUS	migMX	migALL	migUS	migMX	migALL
TTE_{g0}	.0006	.0009	.0015	.0005	.0011	.0017
	(.0015)	(.0023)	(.0030)	(.0011)	(.0015)	(.0021)
TTE_{g1}	.0024	-.0002	.0024	.0011	.0002	.0009
	(.0017)	(.0021)	(.0027)	(.0015)	(.0019)	(.0028)
TTE_{p0}	.0015	.0000	.0016	.0006	.0005	.0012
	(.0013)	(.0020)	(.0026)	(.0009)	(.0014)	(.0020)
TTE_{p1}	.0003	.0016	.0025	-.0003	.0023	.0009
	(.0031)	(.0039)	(.0049)	(.0023)	(.0037)	(.0052)
TTE_{p2}	-.0003	.0050	.0055	.0903***	.0038	.0080
	(.0051)	(.0050)	(.0075)	(.0317)	(.0041)	(.0071)

Standard errors clustered at the village level.

school children (4% of the total 1997 sample) are significantly more likely to have foreign migrant members if they are based in Progresa villages. This corresponds to two families among the group of 334 households in treatment villages having foreign migrants, while no household among the 161 control ones has international migrants. Hence, the significant difference may be due to the small size of the cells and to the very few international trips within cell. In the remaining part of the analysis, particular care will be taken in comparing migration rates for groups that are potentially different in 1997. Finally, I estimate (3.8) using 1997 data: the coefficients β_4 and β_{5j} are never statistically different from zero.[17]

3.5 Aid and migration: results

This section presents estimates of the program impact on various migration measures, both at the individual and at the household level. The advantage of using individual-level data is that the larger sample size increases the

[17]This latter set of results is not shown, but it is available upon request.

precision of the estimates. Its potential drawback is that it may not truly represent the decision-making process, because number and type of migrants may be chosen simultaneously at the household level. Thus, I also estimate the impact of the program components on migration using families as the unit of analysis. I look at the program effect on the likelihood of having at least one migrant in the household using a probit model. I estimate the individual likelihood of being a domestic or international migrant by multinomial logit, the ones at the household level by bivariate probit, and the likelihood of being a migrant irrespective of the destination by probit for both individual- and household-level data.

In order to improve the precision of the estimates, I add the following set of conditioning variables to all regressions. Number of household members aged 14 to 40; size of owned land; dummies for whether household suffered from a series of "shocks" during the interview year; age of household head (or spouse, in case of missing information). This first set of regressors is added to both individual- and household-level regressions. Age (as second-order polynomial), gender and number of domestic and international migrants (excluding self) are added to individual-level specifications only. I also experimented with additional regressors, such as presence of disabled individuals; individual temporary migration experience; village migration intensity; type and number of animals owned; federal state dummies. Including these variables does not change the results and has no sizeable effect on the standard errors. All household-specific variables are from the 1997 survey, excluding shocks.

Robust standard errors are clustered at the village level, since localities are the primary sampling units and the randomization is performed at the village level.

3.5.1 Average program effect on migration

Table 3.6 provides group means of the three measures of migration, for domestic, international and pooled labour migrants. Average U.S. migration

Table 3.6: 1998 and 1999 average migration levels

	Valid 1998 observations			Valid 1999 observations		
	All	US mig	MX mig	All	US mig	MX mig
Treatment						
Individuals	.0455	.0112*	.0342	.0317	0.0078	.0230
Households (% mig>0)	.0324	.0141*	.0192	.0322	.0150*	.0187
Households (mean)^	1.487	1.407	1.471	1.465	1.353	1.459
Control						
Individuals	.0464	.0070*	.0393	.0360	.0097	.0264
Households (% mig>0)	.0239	.0088*	.0153	.0260	.0087*	.0188
Households (mean) ^	1.504	1.552	1.454	1.383	1.486	1.291
P-value of mean difference[a]	0.878	0.065*	0.338	0.283	0.703	0.258
P-value of mean difference[b]	0.107	0.100*	0.364	0.268	0.070*	0.989
P-value of mean difference[c]	0.884	0.389	0.920	0.420	0.441	0.151

^ computes mean migrant number for household with at least one migrant.

[a] p-value of the mean difference at the individual level.

[b,c] p-value of the mean difference at the household level$b$=proportion; c=mean).

is significantly higher for the treatment group both at the individual and at the household level. Average U.S. migration is 1.1% for treatment individuals and 0.7% for control ones. Hence, program availability is associated with a 60 percentage point increase in average migration rate. The proportion of households with at least one international migrant rises from 0.9 to 1.4, corresponding to a 60 percentage point change. However, average household migration does not differ between the two groups.

I interpret this fact as evidence that the program transfers are associated with new households beginning to send their members abroad, rather than households with existing U.S. migrants intensifying their members' migration rates. This may be either because families with pre-program U.S. migrants were either less financially constrained than others to begin with.

Alternatively, the presence of migrants and their remittances may loosen the previously existing financial constraints. This may imply that no migrant "rationing" occurred in households with migrants before the program implementation (I abstract from causality issues here). The difference in proportion of households with international migrants persists in 1999. In November 1999 treatment households have a 0.72 percentage point higher likelihood of having at least one U.S. migrant than control households. However, the difference in individual migration rates is no longer statistically significant.

I find the fact that the program may be associated with higher U.S. migration already few months after its beginning very interesting, because very little money had been transferred at that time. The first grants were distributed in May 1998 and, given that there is no schooling in the summer months, the more substantial cash component of the program - the schooling subsidy - is received (for the first or second time) in October-November. The existence of the program and the certainty of eligibility may have loosened financial constraints for poor households also through general equilibrium effects. The link between program availability and loosened financial constraints may operate through two channels: first, the higher liquidity in the communities brought about by the program's cash injection may have increased credit availability. Second, the stream of certain earnings associated with program eligibility may be used as a collateral to borrow. These issues are not further investigated here because they go beyond the scope of the current analysis. However, they clearly deserve more attention. Domestic migration rates do not differ between the control and treatment group. This fact is consistent with domestic migrations being less costly than international ones. Hence, trips within Mexico may be financed more easily than U.S. ones.

Table 3.7: 1998 individual migration difference for various household types

	Individual			Household		
	migUS	migMX	migALL	migUS	migMX	migALL
TTE_{g0}	.0004	-.0001	.0003	.0022	.0029	.0053
	(.0008)	(.0025)	(.0030)	(.0018)	(.0040)	(.0047)
TTE_{g1}	.0010*	.0026	.0048*	.0039**	-.0045	.0006
	(.0006)	(.0025)	(.0028)	(.0019)	(.0049)	(.0053)
TTE_{p0}	.0005	.0010	.0016	.0021	.0006	.0030
	(.0007)	(.0022)	(.0027)	(.0016)	(.0038)	(.0043)
TTE_{p1}	.0017*	.0051	.0088**	.0050	.0032	.0101
	(.0009)	(.0036)	(.0039)	(.0029)	(.0067)	(.0073)
TTE_{p2}	.0005	-.0042	-.0031	.0063	-.0066	-.0018
	(.0014)	(.0045)	(.0052)	(.0044)	(.0092)	(.0102)

Standard errors clustered at the village level. Tests of the IIA assumption show that odds are independent of other alternatives.

Table 3.8: 1999 migration difference for various household types

	Individual			Household		
	migUS	migMX	migALL	migUS	migMX	migALL
TTE_{g0}	-.0008	-.0030*	-.0048	-.0018	-.0046**	-.0074***
	(.0010)	(.0018)	(.0027)	(.0017)	(.0022)	(.0031)
TTE_{g1}	.0004	.0003	.0011	-.0010	.0016	.0002
	(.0008)	(.0015)	(.0021)	(.0015)	(.0020)	(.0026)
TTE_{p0}	.0006	-.0022	-.0017	-.0007	-.0010	-.0022
	(.0008)	(.0016)	(.0022)	(.0013)	(.0019)	(.0025)
TTE_{p1}	-.0013	.0027	.0014	-.0026	-.0003	-.0016
	(.0011)	(.0024)	(.0031)	(.0037)	(.0033)	(.0044)
TTE_{p2}	-.0014	-.0016	-.0041	-.0032	-.0062	-.0123**
	(.0013)	(.0030)	(.0038)	(.0048)	(.0049)	(.0062)

Standard errors clustered at the village level. Tests of the IIA assumption show that odds are independent of other alternatives.

3.5.2 Heterogeneous treatment effects

Tables 3.7 and 3.8 present the TTEs for households with similar size and type of offspring obtained from (3.6). In this way, one can observe which group is affected by the program.

This exercise reveals that in 1998 treated households with above average grant amount are significantly more likely to have U.S. migrants than similar families in the control group. Further investigation reveals that, although all large transfer recipients have higher migration rates than the control group, the difference is particularly marked among households with a low to medium proportion of conditional grant (i.e. those with $p = 1$). One possible interpretation of these results, consistent with the credit constraint hypothesis, is that families entitled to sufficiently large transfers may use them to fund international trips either directly or as collateral to borrow against (hence the absence of treatment effects for low grant sizes). However, this does not happen when subsidies are linked to the home stay of potential migrants (hence the absence of treatment effect for recipients of large grants due mainly to secondary scholarships). Since the evidence in Table 3.7 is insufficient to disentangle the two effects, I will return to this point when commenting the next set of estimated parameters.

Whichever the driving force behind the significant differences in 1998, they disappear in 1999. According to Table 3.8, in fact, not only large grant recipients stop having different migration rates than control individuals. Also, recipients of low levels subsidies present a significantly lower rate of domestic migration. The next sets of results will shed more light on this issue. To conclude, note that the significant negative sign of TTE_{p2} is consistent with the predicted conditionality effect.

I now proceed to assess the impact of the two program components - conditional and unconditional transfers. I test their impact by regressing the individual and household migration likelihood on controls for potential household grant size and its proportion due to male secondary school attendance, as in (3.8).

Table 3.9: Marginal effects of program components - individual-level 1998 data

	(1)		(2)		(3)	(4)
	Multinomial logit				Logit	
	migUS	migMX	migUS	migMX	migALL	migALL
treatment dummy	.0009	.0000	.0011	.0002	.0043*	.0062
	(.0009)	(.0036)	(.0017)	(.0054)	(.0025)	(.0040)
proportionSM	.0017	-.0054	.0163**	.0298	-.0012	.0182
	(.0019)	(.0064)	(.0086)	(.0227)	(.0054)	(.0170)
proportionSM^2			-.0243*	-.0564*		.0314
			(.0138)	(.0337)		(.0264)
hh grant (x10^4)	-.0079	.0303	-.0107	.0345	-.0401	-.1130
	(.0118)	(.0419)	(.0126)	(.0521)	(.0392)	(.1472)
hh grant^2(x10^4)			-.0564	-.1810		.4290
			(.0692)	(.2038)		(.6731)

The above continuous variables have been interacted by the treatment dummy. Standard errors clustered at the village level. Tests of the IIA assumption show that odds are independent of other alternatives.

Table 3.9 presents the estimates of the parameters of interest when I consider the variables as first- and second-order polynomials, alternatively. The coefficients from the linear specification for both domestic and international migration are not statistically significant. The quadratic specification instead reveals that U.S. migration among the treated group is a positive function of grant composition, for households where less than one third of the transfer is linked to male secondary school attendance (for whom the individual foreign migration likelihood is 0.2 percentage point higher than the control group). Beyond that value the relationship is inverted: the higher the proportion of the grant due to male secondary scholarships in the household, the lower the likelihood of being an international migrant. People from families with the highest proportion of conditional grant are 0.3

percentage points less likely to be abroad. A possible explanation for this specific functional form may be that the grant income effect dominates the composition one initially because the money is mainly transferred unconditionally. When the proportion of conditional grant increases, households shift their offspring's time allocation from migration to schooling.

Table 3.10: Marginal effects of program components - household-level 1998 data

	(1)		(2)		(3)	(4)
	Bivariate Probit[a]				Probit	
	migUS	migMX	migUS	migMX	migALL	migALL
treatment dummy	-.0031	.0106**	.0037	.0107	.0074	.0138
	(.0039)	(.0051)	(.0067)	(.0087)	(.01354)	(.2448)
proportionSM	-.0003	.0196	.0085	.0586	.0193	.0750
	(.0088)	(.0152)	(.0266)	(.0389)	(.0179)	(.0489)
proportionSM^2			-.0138	-.0631		-.0895
			(.0393)	(.0567)		(.0713)
hh grant (x10^4)	.1080*	-.1680**	-.1883	-.1599	-.0457	-.3268
	(.0574)	(.0857)	(.2811)	(.3634)	(.1038)	(.4668)
hh grant^2 (x10^4)			.2253*	-.1997		1.9159
			(.1357)	(.8528)		(3.3612)

The above continuous variables have been interacted by the treatment dummy. Standard errors clustered at the village level. [a]: rho is positive and significant.

The likelihood of being a domestic migrant decreases more than proportionally as the proportion of conditional grant goes up. From Table 3.9, individuals in households with maximum conditional grant are 4 percentage points less likely to migrate than those in families with no eligible secondary school males. Analysis at the household levels reveals that Mexican migration is also a negative function of grant size: 100 extra *pesos* decrease migration by 0.16 percentage points. However, most eligible households have higher domestic migration rates than control ones, as shown by the significance of the treatment dummy coefficient. Only households with a potential grant size exceeding 630 bimonthly *pesos* have an overall decrease in Mexican migration.

Household-based data analysis also reveals a pattern broadly consistent with the predictions of the model, although not necessarily with the evidence from individual-level data. Bivariate probit estimates show a positive correlation between grant size and international migration. The coefficient from the first column of Table 3.10 implies that 100 extra *pesos* increase the likelihood of having international migrants in the treated group by 0.1 percentage points (corresponding to a 7% increase at mean values). The effect is even smaller when using the result from the quadratic specification.

To sum up, the results from 1998 show that Progresa is associated with an increase in international migration coupled with a reduction in domestic ones, to some extent. Both pieces of evidence are consistent with the view that credit constraints may influence the household migration decisions. As not all desired international migrations can be financed, some members will be allocated to domestic migration. The size of the effects is small, though. A possible explanation may be the recent implementation of the program: in November 1998, households had been receiving schooling grants (which constitute the bulk of the financial assistance) only once or perhaps twice. The little money actually received may limit their scope for changes in migration intensity and destination.

The magnitude of the effect in 1999 does not appear to be larger, though, at least for international migration. The individual likelihood of being in the US is 0.23 percentage point lower for those in families where half of the grant is due to male secondary scholarship, and 0.38 percentage point lower for households with the maximum proportion of conditional grant. Moreover, the evidence from Table 3.11 (columns 2 and 4) is to some extent puzzling: for households with conditional transfers exceeding one third of the total, a higher proportion of conditional grant increases the likelihood of having one or more foreign migrants. This issue will be further investigated

Table 3.11: Marginal effects of program components - individual-level 1999 data

	(1)		(2)		(3)	(4)
	Multinomial logit				Logit	
	migUS	migMX	migUS	migMX	migALL	migALL
treatment dummy	-.0003	-.0015	-.0015	.0016	-.0062*	.0001
	(.0014)	(.0025)	(.0026)	(.0039)	(.0039)	(.0071)
proportionSM	-.0046**	.0036	-.0083	.0320**	-.0022	.0363
	(.0023)	(.0047)	(.0065)	(.0152)	(.0071)	(.0235)
proportionSM^2			.0049	-.0454*		-.0601
			(.0081)	(.0239)		(.0308)
hh grant (x10^4)	.0179	-.0010	.0641	-.2486**	.0663*	-.0347
	(.0185)	(.0333)	(.0745)	(.1268)	(.0408)	(.3154)
hh grant^2(x10^4)			.3257	1.6083*		1.8204*
			(.7754)	(.9162)		(1.0002)

The above continuous variables have been interacted by the treatment dummy. Standard errors clustered at the village level. Tests of the IIA assumption show that odds are independent of other alternatives.

below. There is also a negative relationship between grant size and U.S. migration, for low potential grant levels. The relation is instead positive for bimonthly subsidies larger than 660 *pesos* and positive for low conditional grant proportions.

The effect of the program component on domestic migration is similar in sign and magnitude to the 1998 one. There are negative grant composition and size effects, both at the individual and at the household level (the quadratic grant function in Table 3.12, column 2 has a negative first derivative throughout its support). As in 1998, the reduction in Mexican migrations for beneficiaries of large grants is more than compensated by an increase in U.S. trips, resulting in an overall positive relationship between subsidy size and individual migration (Table 3.12, columns 3 and 4).

Table 3.12: Marginal effects of program components - household-level 1999 data

	(1)		(2)		(3)	(4)
	Bivariate Probit[a]				Probit	
	migUS	migMX	migUS	migMX	migALL	migALL
treatment dummy	-.0012	.0048	.0044	-.0057	.0035	-.0006
	(.0046)	(.0055)	(.0084)	(.0123)	(.1375)	(.2790)
proportionSM	-.0016	-.0093	-.0559**	.0653	-.0101	.0045
	(.0033)	(.0124)	(.0299)	(.0402)	(.0147)	(.0489)
proportionSM^2			.0881**	-.1082*		-.0223
			(.0471)	(.0605)		(.0732)
hh grant (x10^4)	.0776	-.0191	-.4910*	.4137	.0712	.2416
	(.0448)	(.0734)	(.2974)	(.4401)	(.1031)	(.5752)
hh grant^2 (x10^8)			3.6808**	-3.4938		-.0922
			(1.8402)	(4.5971)		(.3546)

The above continuous variables have been interacted by the treatment dummy. Standard errors clustered at the village level. [a]: rho is positive and significant.

To conclude, the evidence provided so far is consistent with the picture emerged from the theoretical model: increasing the income of financially constrained household may result in an increase of costly migration, which could not be funded before. This is shown both directly (significantly larger number of U.S. migrants) and indirectly (through the possible substitution of domestic for international trips). The magnitude of these effects appears to be quite small.

The current analysis presents two main shortcomings: first, it is hard to reconcile the positive relationship between the composition effect and 1999 household U.S. migration with the theoretical predictions. Second, the two program components are never jointly significant in a single specification. This might be caused by the limited variation in the data caused by the few observed migrations. In order to test for the joint significance of the effect, and to test for the robustness of the results, I relax the functional form assumption and proceed with the estimation of cell means, as explained above.

Table 3.13: Marginal effects of program components - 1998

	(1)			(2)		
	1998					
	Individual-level data			Household-level data		
	migUS	migMX	migALL	migUS	migMX	migALL
TTE_{g10}	.0000	.0027	.0033	.0124**	-.0088**	-.0009
	(.0012)	(.0038)	(.0044)	(.0077)	(.0044)	(.0076)
TTE_{p10}	.0015	.0024	.0055	.0025	.0174	.0205
	(.0021)	(.0052)	(.0062)	(.0079)	(.0158)	(.0174)
TTE_{p20}	.0000	-.0054	-.0054	-.0020	.0085	.0054
	(.0016)	(.0032)	(.0041)	(.0047)	(.0134)	(.0132)
TTE_{p21}	-.0010	-.0071	-.0093*	-.0036	-.0047	-.0091
	(.0012)	(.0038)	(.0045)	(.0046)	(.0069)	(.0083)

Tests of the IIA assumption show that odds are independent of other alternatives. The correlation between the residuals from the household-level estimation is positive and significant. Standard errors clustered at the village level. Regression estimates available upon request.

The TTEs presented in Tables 3.13 and 3.14 are estimated from (univariate or bivariate) probits of the likelihood of being a migrant, or of having migrants in the household, interacting the treatment dummy by the discrete variables grouping grant level and composition, as described above. The estimated effects suggest that both domestic and U.S. migrations are a function of grant size and composition. The effects tend to be more marked in 1999, when more money has been distributed and households have had enough time to respond to the new set of incentives.

As regards international migration, the effect of the program on migration is larger for high grant sizes, i.e. the likelihood of having migrants in

Table 3.14: Marginal effects of program components - 1999

	(3)			(4)		
			1999			
	Individual-level data			Household-level data		
	migUS	migMX	migALL	migUS	migMX	migALL
TTE_{g10}	.0039**	.0020	.0073*	.0243***	-.0096*	.0030
	(.0023)	(.0026)	(.0043)	(.0103)	(.0046)	(.0088)
TTE_{p10}	-.0023***	.0045	-.0010	-.0086***	.0155	-.0015
	(.0006)	(.0045)	(.0042)	(.0020)	(.0157)	(.0109)
TTE_{p20}	-.0020***	-.0000	-.0042	-.0000	-.0046	-.0030
	(.0005)	(.0032)	(.0031)	(.0053)	(.0055)	(.0084)
TTE_{p21}	.0006	-.0033	-.0034	.0248*	-.0112**	-.0016
	(.0016)	(.0026)	(.0037)	(.0204)	(.0039)	(.0116)

Tests of the IIA assumption show that odds are independent of other alternatives. The correlation between the residuals from the household-level estimation is positive and significant. Standard errors clustered at the village level. Regression estimates available upon request.

the household is a positive function of transferred resources. A household that receives a high grant level is (in 1998 and 1999, respectively) 1.2 and 2.4 percentage points more likely to have at least one US migrant member than one that receives a low grant level (when both are compared with the respective control group). The same applies to individual data in 1999. This is consistent with the credit constraint hypothesis.

Conditioning the grant to secondary school attendance reduces contemporaneous U.S. migration. In 1999, individuals from households with at least one secondary school eligible male are 0.2 percentage point less likely to be U.S. migrants than those in households where all transfers are unconditional. However, households with a high proportion of conditional transfers are more likely to have spells of U.S. migration than those with low to medium conditional grants. A similar conclusion had been reached when imposing explicit functional forms. A possible explanation of this migration pattern will be pursued below.

The relationship between domestic migration and grant size is negative. This fact may be interpreted in two different ways. Households may be shifting migrants from Mexico to the US, as they are able to fund the more expensive, yet potentially more rewarding international trips. Alternatively, this result may be evidence of the "home bias" effect: the additional income received through Progresa makes migration more costly (for instance because of the higher level of forgone home consumption), resulting in fewer members being away. The first interpretation seems more sensible. If the "home bias" effect were the dominating one, there would not be a surge in international migration.

As far as the conditionality effect is concerned, there is evidence that higher proportions of conditional grants are associated with lower migration rates also for Mexican migration in 1999.

To test the robustness of the result, I tried to condition on different sets of households. Moreover, I included measures of the conditional grant proportion associated with female secondary scholarships. The results did not change in either case.

Effect of a cap on maximum transfer size on international migration

The previous analysis has revealed that households with a high proportion of conditional transfers are more likely to have spells of U.S. migration in 1999 than those with low to medium conditional grants. Here I check whether this may be caused by the subsidy cap. The maximum bimonthly amount that households are entitled to is 1170 *pesos* in November 1998, and 1390 the following year. Households with a number of eligible children that exceeds the maximum subsidy have the individual school grants re-scaled to sum up to the maximum level. Hence, their monetary incentives to send members to school are lower, and the comparative incentives for migration higher.

Visual observation of cell means confirms this intuition: mean U.S. migration for treated households - but not for control ones - decreases when capped households are excluded from the computation both in 1998 and 1999.

To test this hypothesis, I first interact the treatment dummy with a variable taking the value of one for those households that are capped because they have an "excess" of secondary school males. I repeat the same exercise with a dummy for all "capped" households, irrespective of their school children composition. The magnitude and significance of the estimated TTEs do not change in either occasion, though.

However, when I create a third grant size category to group capped households (hence the new g takes three values, 0, for up to medium potential grant size, 1, for higher than average grant size, 2, for maximum grant size and an "excess supply" of eligible school children), the significance of TTE_{p21} becomes much weaker (with a p value of .133), and its magnitude slightly smaller, dropping to 1.9. This only holds when considering the households with "too many" eligible secondary school males, and not with all capped households. I interpret this fact as evidence that the presence of a cap on the maximum size of the transfer may increase migration (although other factors may be present too). This result confirms once more how the different modes and formats of aid to developing countries may have different impacts on household time allocation, even when the same amount of money is being transferred.

3.5.3 Medium-term migration

So far I have shown that Progresa is associated with a short-term increase in average international migration. I have inferred that this is due primarily to the existence of an unconditional transfer component, i.e. that households are transferred money without having to comply with requirements that force them not to migrate. However, I have also shown that conditional transfers targeting prospective migrants may achieve a migration reduction in the short run. In fact, households with a sufficiently high proportions of secondary school males may end up having lower migration rates than before the program implementation because of its "conditionality" effect (i.e. the necessity to comply with the program requirements in order to be eligible for the subsidy). However, it is possible that secondary school males may revert to migration after they complete secondary education. The model shows that their future choices depend on the location where their accumulated human capital reaps the highest return - and that such location may not necessarily be abroad, but rather in Mexican urban centres such as state capital.

One way to assess the medium-term program effect on migration is to look at choices of individuals who have completed the first three years of secondary school. I focus on the sub-set of individuals aged 15 to 18 with some or complete junior high-school education in November 1999. These individuals were potentially entitled to the program educational grant for the 1998-1999 academic year. I then look at whether children from treatment villages are significantly more likely to migrate than children from the control group.

I create two dichotomous variables that record migrations for two categories of teenagers: those with complete junior high school in 1999, and those with some level of secondary education. I focus on trips started in the six months between the end of the academic year, which I assume to be at the beginning of June, and the interview time, in November. The valid sample consists of 1567 individuals with completed junior high school, and of 3602 ones with at least one year of secondary education.

Mean comparison of the individual likelihood of undertaking a migration shows that children from treatment localities do not have different migration propensities than those in control ones. Although the size of the sample is quite small, and the fact they do not migrate immediately after the end of the transfer does not prevent them from undertaking future migration, this result suggest that education support programs may not cause higher migration in the medium term. One shortcoming of the current analysis is

Table 3.15: Average migration rate of 1999 secondary schoolchildren

Sub-sample	Complete junior high	$\geq$1 year secondary education
Treatment	.024	.013
	(.007)	(.003)
Control	.023	.014
	(.006)	(.004)
p-value of difference	.933	.776

Standard errors clustered at the village level.

that ideally one would want to observe post-schooling migration behaviour of individuals who began to go to junior high school because of the program implementation, rather than those who took advantage of the scholarship while they had already started the schooling cycle. Unfortunately, the first such cohort graduates in 2001. In 1999 only the latter group is observed. More research is needed in this field.

3.6 Conclusions

The current exercise contributes to our knowledge of the relationship between aid and migration by analysing the impact of the development program Progresa on domestic and international migration of poor rural households. In particular, I have tried to understand how different program components may provide opposite incentives to migrate. The theoretical model has shown that unconditional transfers may increase migration rates by loosening the credit constraints faced by indigent households. The larger the poverty level, the higher the likelihood that this effect dominates any (supposed) "home bias", often described in the literature as a preference for home consumption. Conditional transfers that target potential migrants and require them to stay at home are associated with a migration reduction in the short run. Individual behaviour in the medium run and their migration choices depend on differentials in returns to the extra skills learned.

The empirical analysis, which considers migrations both as an individual and as a household decision, uncovers interesting facts. First, the program is associated with an increase in average international migration, confirming that fact that credit-constrained households may be forced to undertake a sub-optimal number of potentially profitable migrations because of the impossibility to finance the trips. Second, the extra cash is used to finance migrations in households who did not have any migrant previously, rather than increase average migration rates for all households. Third, average domestic migration levels are not affected by program availability, suggesting that it is mainly the more costly international trips that cannot be financed by poor households.

Fourth, the existence of the program may increase the level of borrowing among poor households, by injecting cash in the villages or by being used as a collateral to access the credit market. This is deduced by observing that international migration rates increase after very little money is transferred to households, and do not grow in a way correlated with the total amount of cash transferred over time (i.e. total program effect does not differ substantially for treatment poor between November 1998 and November 1999).

Fifth, different households face varying migration incentives linked to their demographic composition. Those with a large grant proportion coming from secondary school subsidies will migrate less than those for whom the program requirements are not binding (i.e. families with primary school children). One exception to this pattern is provided by 1999 U.S. trips (when modelled as decided at the household level, rather than at the individual one), which appear to be higher for high than for medium conditional grant proportions. A partial explanation is given by showing that households where not all secondary school males are entitled to the school subsidy because of a cap on the maximum grant size are more likely to have U.S. migrants. However, additional (and yet uncovered) factors are likely to play a significant role in explaining this puzzling effect.

Sixth, as (potential) grant size increases, households substitute (cheaper but with lower return) domestic for (costlier but with a higher benefit) international migration. This fact provides additional evidence of the importance of credit constraints in developing countries. Nevertheless, the net creation of migration associated with Progresa is positive.

Seventh, the secondary school subsidy is not associated with a post-program increase in migration. I observed the location choice of teenagers with completed junior high school in 1999. There is no significant post-program different migration rate between individuals from treatment and control groups.

These findings have interesting policy implications. They show that it is possible to implement aid policies that do not result into higher international migration. The way to insure that the transfers are used to finance extra migrations is to give emphasis to conditional grants such as subsidies for secondary education. This approach is already being pursued, albeit perhaps unintentionally. In fact, Progresa has been recently extended to cover Mexico's urban areas (under the name of Oportunidades). There, the education transfer has been extended in order to subsidize also additional schooling years. It is possible that this type of transfer may result into a net reduction in short-term international migration. If skilled wages in urban Mexico are sufficiently high, the program may not even cause higher medium-term U.S. migration.

The current research has also pointed out directions for further investigation. One is to understand the effect of the program on access to credit. The other is to look at comparative effects of different policies on international migration, such as aid and border enforcement.

3.7 Appendix: variables creation

In this section, further information is provided on the creation of the variables of interest, potential grant size and composition.

No information is available on effective size of the received transfer. Hence, it is only possible to compute the potential transfer size. This approach rules out endogeneity issues related to partial acceptance of the program, while it permits the estimation of parameters with strong policy relevance: the impact of the availability of a transfer (the policy maker cannot force individuals to comply with the program requirements). Its drawback is the difficulty of assessing the effective monetary impact of the program.

Potential grant size is computed considering all children aged 5 to 18[18] in the November 1998 survey using last completed school grade. I use 1998 information rather than 1997 one because I would like to include current migrants in the computed potential grant, given that some of them, especially those who are in the same state, may actually decide to go back to school. No information on schooling level of migrants is available in the September 1997 survey. A consequence of this choice is that there will be differences in 1999 potential grant size estimated using 1997 versus 1998 data. In the latter case, only children who completed the first year of primary school in the 1997-1998 academic year will be considered eligible for the program in the academic year 1999-2000 (having assumed that they completed their second grade in 1998-1999). Potential grant size built projecting grade completion information for the 1996-1997 academic year, instead, must assume that all children aged 6 in September 1997 will have completed second grade by June 1999. Only if I match grade completion data with current enrolment for first-graders will the two measures coincide in 1999. A further difference is given by the fact that using 1998 information I do not have to assume that nobody fails in the 1997-1998 academic year.

Two implicit assumptions are made in the computation of potential grant size: first, grade completion is independent from existence of the program. The assumption is violated if the latter is a function of future program eligibility. Children progressing to a subsidized grade may have higher incentives to pass, while those in the final subsidized grade may be more likely to fail to receive benefits the following year. Second, it is assumed that all students

[18] 18 year old individuals are dropped from the computation in 1999. Children aged 5 in November 1998 are included in case they reach third grade in the 1999-2000 academic year.

pass the grade, independently of actual school attendance and of whether the child actually left the household. Again, any strategic behaviour has been hence ruled out.

Four measures of potential grant amount are built and used in the estimations. Such measures are obtained by varying two parameters: the type of children included and the time dimension considered. As regards the former, in one case only November 1998 residents are included in the computation, hence omitting all children who are elsewhere for any sort of reason. The time dimension varies in the sense that in one case I estimate current bimonthly transfer size computed for both November 1998 and November 1999, while in another case I compute total potential amount. In this way, emphasis is given to the fact that the cumulative grant, rather than the current one, matters for financially constrained households. There is some variation among the two only in 1999, given by families with children in the third year of junior high in the 1998-1999 academic year.

The figures shown below are computed excluding migrants. Moreover, for 1999 data, they only use the estimated value of potential current grant, instead of using the cumulative one. However, the results are robust to the use of the alternative measures of potential grant and its composition.

Table 3.16: Average pre-program potential grant size (at 1999 values) and grant composition

Variable	Grant size	Composition[a]
Treatment	560	.055
	(8.2)	(.003)
Control	562	.050
	(10.3)	(.003)
p-value of difference	0.906	0.286

[a]: grant composition measured as proportion of grant due to male secondary school attendance.

Standard errors clustered at the village level.

Identification of the parameters of interest in the assessment of the effect of the program components is possible for the following reasons. First, the number and composition of eligible schoolchildren by households is very similar among control and treatment villages and it insures that potential grant size has a common support for the two groups. I created potential grant size and grant composition measures also for the 1997 sample, using 1999 grant size.[19] There is no significant difference for both potential grant size and grant composition between control and treatment households. Second, there is a large variation in the number, gender and school level of eligible children by family also within treated households. Indeed the correlation between grant size and composition is positive and significant, amounting to 0.47 (0.46 for treatment poor only) in 1997, but it is far from unity. Third, the specific grant amounts are such that different combinations of primary and secondary school beneficiaries yield the same potential transfer level. Hence, households entitled to roughly equivalent grant levels may vary substantially in the proportion of secondary school males.

[19] I.e. the size and composition of grant that one household would be entitled to in 1997, had the program been already implemented with 1999 scholarship levels.

Chapter 4

A note on the identification of counterfactuals in the experimental sample of Progresa

4.1 Introduction

Progresa targets Mexican poor rural households and provides grants to improve education, health, consumption and the role of women in the household. Its main monetary component, apart from a smaller nutritional subsidy, is in the form of a schooling subsidy to children attending the last third grades of primary school and the three grades of secondary school. Transfers are made to women only (normally the spouse of the household head), and are conditional upon regular visits to health centres, to *platicas* where women are taught about health and nutrition issues, and to a school attendance rate of at least 85% of term time. 186 of the 506 villages sampled for evaluation purposes are randomized out. Poor residents of these villages are not administered the program until 2000. Households are classified into poor and non-poor according to the information collected in the pre-program

September 1997 census of Progresa localities. All residents of both control and treatment villages are then interviewed at biannual intervals. Detailed data are collected on health, consumption, income and employment, education and migration at least in one of the two annual surveys. With a sample size ranging between 22,000 and 25,000 households in both control and treatment villages, complete coverage of all locality residents, a panel of up to five waves, of which one or two prior to the implementation of the program, and the exogenous variation induced by the randomization, it is not surprising that the Progresa data have attracted the attention of scholars and researchers.

One shortcoming associated with the implementation of Progresa is the presence of a group of eligible households who did not receive any benefit by March 2000. Buddelmeyer and Skoufias (2003) report that this group amounts to 27% of the total eligible population in treatment localities. The likely explanation for this fact is attributed to administrative delays and mis-classifications. Albarran and Attanasio (2001) compare forgotten and beneficiary households, and find that there are significant differences in average observable characteristics, although no set of variables could identify either group with probability one.[1] The absence of an observable counterfactual for eligible households who did not receive the transfer complicates the estimation of the program's impact on actually treated subjects.

This issue has been seldom recognized and addressed appropriately by previous users of the Progresa data. In most cases, this estimation issue is ignored altogether. Schultz (2004) and Dubois et al. (2001), among others, do not mention it at all. Buddelmeyer and Skoufias (2003) use only a sub-sample of the available eligible poor. Albarran and Attanasio (2001) introduce additional assumptions to identify the effect of the treatment on the eligible non-treated.

This note reviews the approaches employed to estimate Treatment on the Treated Effects ($TTEs$) in the presence of non-compliers and presents

[1] Albarran and Attanasio (2001), p.9.

an alternative estimating procedure to identify counterfactuals that requires no identifying hypotheses nor exclusion restrictions for users of the Progresa data. Although the focus of the current analysis is on mean effects only, the same logic can be applied to distributions.

4.2 Standard estimation of $TTEs$ with non-compliers or dropouts

Suppose one wishes to estimate the impact of Progresa on an outcome y. The program's design is such that a random subset of the whole sample of participants is not administered the treatment, named control group in jargon. The associated mean outcomes are $E[y|P = 1]$ and $E[y|P = 0]$ for poor households in the treatment ($P = 1$) and control ($P = 0$) groups, respectively. The randomization insures that the distribution of observable and unobservable characteristics among treatment and control subjects is the same. In such case, the control group is a valid counterfactual for the treatment one, i.e. if the treated agents had not received the program, their average outcome would have been $E[y|P = 0]$. Hence, the treatment on the treated effect is simply the difference of mean outcomes among the two groups,

$$E[y|P = 1] - E[y|P = 0] \tag{4.1}$$

Assume instead that a non-random sub-sample of the treatment group of size n_t does not actually receive the program. In consequence, the effect of the program on potential and actual beneficiaries may differ and (4.1) would simply measure the intention to treat. Mean outcome for agents in the treatment group would be a weighted average of mean outcomes for actually treated and non-treated agents,

$$E[y|P = 1] = (1 - n_t)E[y|P = 1, C = b] + n_tE[y|P = 1, C = nc] \tag{4.2}$$

$$E[y|P=0] = (1-n_c)E[y|P=0, C=b] + n_c E[y|P=0, C=nc] \quad (4.3)$$

where C equals b for *de facto* beneficiaries and nc for non-complying households and $n_c = n_t = n$ because of the randomization. In order to estimate the effect of the program on actual beneficiaries only, one would have to know who would be actually treated among control subjects. Since $E[y|P = 0, C = b]$ is not observed, the impact of the program on actually treated participants cannot be directly estimated.

Note that the described scenario is quite common in social experiments, as some individuals from the treated group may fail to comply with the eligibility requirements or may drop out of the program. In these circumstances estimation of the parameter of interest has to rely either on some identifying assumption or on some exclusion restriction. A standard identifying assumption employed by the literature in similar cases is that the program has no effect on non-beneficiaries in the treatment group:

$$E[y|P=0, C=nc] = E[y|P=0, C=nc] \quad (4.4)$$

This implies that program impact for actual beneficiaries is a proportion of the difference in mean outcomes between treatment and control groups (irrespective of actual treatment)

$$\frac{E[y|P=1] - E[y|P=0]}{n} \quad (4.5)$$

as obtained by rearranging (4.2) and (4.3) after imposing (4.4). This is the assumption used by Bloom (1984). Heckman *et al.* (1998) explore the use of exclusion restrictions to identify the parameter of interest when the impact on non-compliers or dropouts is not expected to be null. However, they show how the results are sensitive to small changes in the underlying assumptions. More recently, Albarran and Attanasio (2001) rely on a weaker assumption in their study of the effect of Progresa on private transfers among poor households. They hypothesize that the effect of the program on eligible subjects that ended up not being treated is equivalent to the general equilibrium effect on the non-poor population.[2] Hence,

$$\begin{aligned} E[y|P=1, C=nc] - E[y|P=0, C=nc] = \\ E[y|NP=1, C=nc] - E[y|NP=0, C=nc] \end{aligned} \tag{4.6}$$

where NP refers to poor and non-poor agents. In this way, the effect of treatment on actually treated households is given by:

$$\frac{E[y|P=1] - E[y|P=0]}{n} - \frac{1-n}{n}\{E[y|NP=1, C=nc] - E[y|NP=0, C=nc]\}$$

4.3 Estimation of $TTEs$ in Progresa

Progresa data permit to estimate $TTEs$ for all eligible poor without the need of identifying assumptions or exclusion restrictions. The estimable parameters bear potential economic and policy relevance.

In order to illustrate the point formally, it is necessary to briefly review the selection process of eligible households employed for the Progresa evaluation and to introduce some notation. The primary source of information are the two reports by Skoufias *et al.* (1999a, 1999b), to which the interested reader is addressed for further details.

In the first stages of the program implementation, in 1997, the selection method identified roughly 52 percent of all households (in both treatment and control villages) as eligible and the remaining 48 percent as non-poor. Let's call the former group "1997 poor". The following year, a revision of the eligibility status was undertaken, since the previous method was felt to discriminate poor households with no young children. A group of households composed by individuals initially classified as not poor was included in the eligible set. This group accounts for a further 25 percent of the total population and it roughly halves the non-poor size. I label them "*densificados*", after the Spanish word used to indicate the eligibility revision

[2] Progresa targets poor households only. However, information is collected also on non-poor ones.

(*densificacion*). A further set of households starts being recorded in the October/November 1998 survey in both treatment and controls villages. They belong to both poor and non-poor group. They are labeled "immigrants". Further details on each group will be provided in the relevant subsections.

Table 4.1: Household categories by village type (absolute values and percentages), November 1998 data

	treatment		control		
	non-poor	poor	non-poor	poor	Total
1997	3233	7837	2048	4682	17800
	(92.4)	(63.6)	(92.6)	(61.4)	(69.4)
densificados	0	3786	0	2491	6277
	(0)	(30.7)	(0)	(32.6)	(24.4)
immigrants	254	667	155	437	1514
	(7.2)	(5.4)	(7.0)	(5.7)	(5.9)
total	3487	12290	2203	7610	25590
	(100)	(100)	(100)	(100)	(100)

Table 4.1 presents a breakdown of households by the aforementioned groups for the 1998 data (round 3). Among poor households, the majority belongs to the "1997" group (63 percent in treatment villages, and 61 percent in control communities). *Densificados* amount to less than a third of all poor, while immigrants constitute 5-6 percent of the sample. At this stage of the analysis it is not clear whether immigrants are part of the sample of eligible individuals, because of their late appearance in the data sets. Assuming they do, outcomes for the treatment and control groups can be decomposed

in the following way[3]:

$$E[y|P=t] = p_{97}E[y|P97=t] \tag{4.7}$$
$$+p_D E[y|PD=t] + (1-p_{97}-p_D)E[y|PI=t]$$

where $t = \{0, 1\}$ indicates control and treatment group, respectively, and $P97$, PD and PI refers to 1997 poor, *densificados* and immigrants. The shares of $P97$ and PD, p_{97} and p_D, are of equal size among treatment and control villages because of the randomisation. The proportion of immigrants in the two types of villages is identical, as shown below.

$E[y|P = 1] - E[y|P = 0]$ estimates the intention to treat effect, i.e. the average effect of Progresa on the eligible sample, irrespective of actual treatment. This parameter is of obvious policy relevance, as it measures the average effect that the program has on the targeted subjects, irrespective of the latter's compliance, which in most cases cannot be directly controlled by the policy maker. However, estimating TT parameters is of particular relevance to understand the impact of Progresa. This is because it turns out that the main source of difference between being a potential and actual beneficiary is due to organisational errors among the program administrators, hence directly controllable by the policy maker. Indeed, households who received zero transfer in spite of their eligibility are actually composed by two different categories: true non-compliers, who chose not to participate in the program, and "forgotten" households, who did not receive the transfer because of administrative errors. The latter group is much larger than the former. This suggests that, had these mistakes not occurred, the average effect of the program may have been different than it actually was.

Table 4.2 shows a breakdown of poor households in Progresa communities by group and compliance status. It reveals that only 73 percent of poor households in Progresa localities are actual beneficiaries.[4] These are

[3]In case they did not, the last object in (4.7) would disappear and the other two terms would be divided by $(p_{97} + p_D)$.

[4]Beneficiary households are all those families in which at least one member receives the program. Non-compliers are households who choose not to participate to the program.

Table 4.2: Poor households in treated villages by treatment status, November 1998 data

	97 poor	Densificados	Immigrants
Beneficiaries	7470	1547	0
	(95.3)	(40.8)	(0.0)
Forgotten	7	2191	688
	(0.0)	(57.8)	(100)
Non-compliers	360	48	0
	(4.5)	(1.2)	(0.0)
Total	7837	3786	688
	(100)	(100)	(100)

Percentages in parentheses. Percentages may not add up to 100 because of rounding errors.

the households one needs to consider in order to estimate TT effects. The parameter of interest, given by $E[y|P=1,C=b]-E[y|P=0,C=b]$, can be decomposed in the following way:

$$E[y|P=t,C=b]= \tag{4.8}$$

$$\frac{p_{97}b_{97}}{p_{97}b_{97}+p_Db_D}E[y|P97=t,C=b]+\frac{p_Db_D}{p_{97}b_{97}+p_Db_D}E[y|PD=t,C=b]$$

as there are no beneficiaries among immigrants. b refers to being *de facto* treated, and b_{97} and b_D are the proportions of household who receive the treatment. The estimation of TT effects in this case is hampered by the difficulty of identifying a valid counterfactual among the control group, as one has to identify the subset of households in the control villages who would have been actually treated, had they been in the treatment group. The next sections deal with this issue, but first I will discuss the outcome from Table 4.2.

All remaining eligible households are classified as "forgotten".

The proportion of forgotten households is not constant across the three groups of 1997 poor, *densificados* and immigrants. As regards true non-compliers, only 360 1997 poor and 48 *densificado* eligible households decided not to participate to the program for unknown reasons. The bulk of zero-transfer households belongs to the "forgotten" group, and that they are mainly concentrated among *densificados* and immigrants[5]. It appears that by March 2000, when the payment records were released, they had not received the benefits, and it is not clear how much they eventually received and when the transfers did actually take place (Skoufias and Hoddinott, 2000). All these individuals with revised status were informed of their changed eligibility, though never entered in the program database (Coady, 2001). Uncovering the different proportion of forgotten households by poverty groups is a first step towards the identification of valid a counterfactual for this category.

The division of zero-transfer recipients into two different categories suggests that only in one case, and for a small set of households, the source of non-compliance is endogenous, i.e. individuals choose not to participate to the program. It is necessary to introduce some *ad hoc* assumption or to find some exclusion restriction in order to identify the program effect for this group.

For the majority of eligible individuals who did not receive the transfer, instead, *de facto* reception of the benefit is exogenous and is irrespective of compliance with program requirements. This suggests that one of the following cases must be true: being forgotten happens on a random basis; being forgotten happens on a non-random basis and one or more variables that predict its likelihood with probability one; being forgotten happens on a non-random basis and follows some criterion that permits to identify a valid counterfactual among the control group; being forgotten happens on a

[5]Again, it is not clear whether immigrants were incorporated in the program at all. The fact that none of the poor ones received the treatment is further support for the conjecture that they do not belong to the eligible group. Note that if immigrants are not incorporated in the program, then they do not belong to the forgotten row of Table 4.2

non-random basis, but there is no criterion that permits to identify a valid counterfactual.

Albarran and Attanasio (2001) thought that the latter case occurred, before distinguishing among the aforementioned groups (1997 poor, *densificados* and immigrants) and categories (beneficiary, non-complier, forgotten). I will show below a criterion to identify counterfactuals for forgotten households. The following sections discuss identification of TT effects for the various subgroups that constitute the sample of eligible households. The details regarding the identification of a counterfactual for forgotten households will be discussed in the densificado section. The assumptions required for the estimation of the program effect for non-compliers will be discussed in a separate section.

The next subsections discuss the hypotheses required for the identification of counterfactuals for each group.

4.3.1 1997 poor

Forgotten households are hardly present among 1997 poor: Table 2 shows that only 7 out of 7837 "1997" poor families are not treated[6], amounting to 0.08 percent of the total. Because of the negligible size of non-compliers, the average effect of Progresa on 1997 poor simplifies to:

$$\begin{aligned}
&E[y|P97 = t] = f_{97}E[y|P97 = t, C = f] \\
&+ n_{97}E[y|P97 = t, C = n] + (1 - f_{97} - n_{97})E[y|P97 = t, C = b] \\
&\cong n_{97}E[y|P97 = t, C = n] + (1 - n_{97})E[y|P97 = t, C = b]
\end{aligned}$$

$C = \{f, n, b\}$ indicates the category the household belongs to (forgotten, non-complier, actual beneficiary), while f_{97} and n_{97} refer to the proportion of households in the first two categories. This parameter measures the program's impact on the poorest treated households. $n_{97} = 0.0008$, and the corresponding 7 households can be simply dropped from the valid sample.

[6]These forgotten households live in four different localities in the state of Guerrero.

Hence, the effect of the program on actual beneficiaries is given by

$$E[y|P97 = t, C = b] = \frac{1}{1 - n_{97}} E[y|P97 = t] - \frac{n_{97}}{1 - n_{97}} E[y|P97 = t, C = n]$$

Some identifying assumptions or valid exclusion restrictions are necessary to estimate the effect of Progresa on non-compliers. These are discussed below.

4.3.2 *Densificado* poor

As concerns *densificados*, only 1.2 percent are true non-compliers, corresponding to 48 families. If one were willing to discard non-compliers because of their small sample size, the only identification issue would regard the presence of forgotten households, who account for 57.8 percent of *densificados.* If forgotten households were a random sub-sample of such group, one could use the whole set of *densificados* in control villages as a valid counterfactual, since

$$E[y|PD = 0, C = b] = E[y|PD = 0, C = f] \cong E[y|PD = 0]$$

Again, this approximation is possible because the proportion of non-compliers is so small to be supposedly negligible. However, the comparison of the distribution of a large group of variables from the pre-program 1997 wave (round 1) shows that there are significant differences in key variables such as wealth, employment and earnings, education and schooling. These results are reported in the Appendix.

By looking at village of residence of forgotten densificados, it appears that entire communities where missed out: all households with revised eligibility in 115 localities out of 302 failed to receive the treatment. Hence, village of residence predicts actual treatment for *densificados* with probability one.

Nevertheless, villages are not missing at random: forgotten localities have on average a significantly lower level of the marginalization index. This index summarises the degree of poverty of each locality, with higher values corresponding to more severe poverty levels. It was computed with the purpose of identifying eligible villages.[7] The Kolmogorov-Smirnov test confirms that the distribution of the index at the village level is not equal between forgotten and treated *densificados*.[8] Indeed, nearly 68% of forgotten localities have a marginalization index lower than the smallest (-0.093) among beneficiary villages, hence forgotten households belong to the least poor villages, as can be noticed in Figure 4.1.

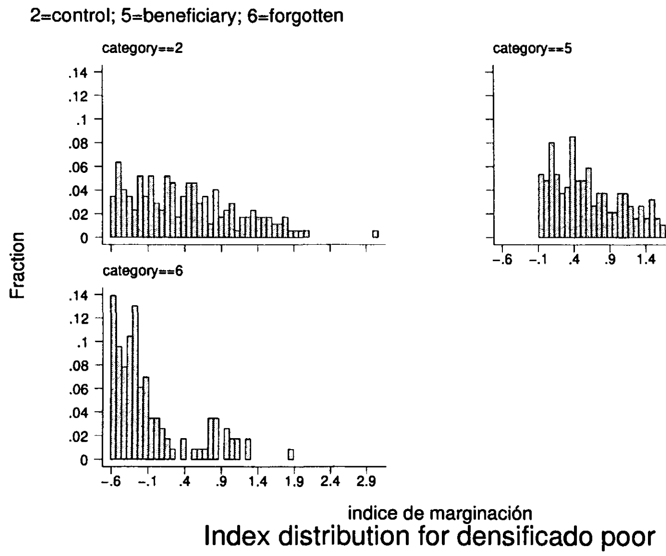

Figure 4.1: *Densificados* - distribution of village marginalization index for control, treated beneficiary and forgotten households

I tested whether, conditional on the marginalization index, no other sets of variables differ significantly between forgotten households and actual beneficiaries, but this hypothesis is not supported by the data.[9]

[7]Further details on the creation of the marginality index can be found in Skoufias *et al.* (1999a).

[8]Whereas the null hypothesis of distribution equality cannot be rejected for all treatment and control villages.

[9]I estimated the likelihood of being forgotten as a function of the household-specific variables contained in the Appendix tables, additionally conditioning on marginalization index expressed as a linear and third-order polynomial, alternatively. I also grouped villages with similar marginalization levels by creating sets of dummies (no villages have

One possible approach to identify valid control villages for treated *densificados* is to group them by the village marginalization level. Naming this index M and setting its realization $m = -0.093$

$$
\begin{aligned}
&E[y|PD = t, C = b] \\
= &\Pr(M > m|PD = t, C = b)E[y|PD = t, C = b, M > m] \\
+ &\Pr(M \leq m|PD = t, C = b)E[y|PD = t, C = b, M \leq m] \\
= &E[y|PD = t, C = b, M > m]
\end{aligned}
$$

i.e. one can use control villages with marginalization index larger than -0.09 as counterfactual for actual beneficiaries. The hypothesis of equal distribution of the village-specific index cannot be rejected, when only villages above the threshold are compared.[10] If one is willing to ignore non-compliers, the above equation permits to estimate the second object to the right of the equal sign in (4.8).

In the same fashion, it is possible to estimate the effect of the program on forgotten households. In this case, though, $\Pr(M \leq m|PD = t, C = f) = 0.68$. While the Kolmogorov-Smirnov test is significant when the two groups of households below and above the m threshold are pooled, it cannot be rejected when considered separately, both for treated and forgotten families.

The TT approach suggested here relies on the village randomisation. Its main shortcoming is that the randomisation is designed for a larger sample of villages, so there is a weaker case for its being valid also for the relevant subset of villages. For this reason, I compare the distribution of variables from the 1997 survey and from the March 1998 consumption module using

the same index level). The remaining variables are either individually or jointly significant.

[10]Note, however, that when we compare the individual poverty level computed by Progresa staff to distinguish poor from non poor, the equality of distribution between treatment and control households is sometimes not rejected. Localities are grouped into seven regions of different size, and different levels of the household poverty index were used as poverty threshold. I compare the distribution of the household poverty index at the regional level because *densificado* and forgotten households are not distributed homogeneously across them.

the groupings described. Overall, it appears that the hypothesis of random program allocation within eligible *densificados* is supported by the data, once households are matched using the marginalization index. The average rejection rate from the analyzed variables is smaller than both the one obtained by comparing intended treatment and control households[11], as in Behrman and Todd (1999), and the one from "unmatched" *densificados* (i.e. comparing all actual treated with all control households). All relevant tables are in the Appendix.

The distribution of the marginalization index among Progresa villages is such that communities with an index lower than -0.093 are forgotten with probability 1, while localities with an index higher than -0.093 have a positive but less than one probability of being forgotten. The local effect of the program around the marginalization index threshold may be estimated using a fuzzy regression discontinuity design (RDD). However, while the quasi-experimental nature of this estimation method may insure that villages around the discontinuity do not differ systematically, nothing guarantees that households in those communities do not differ both in observable and unobservable characteristics. No attempt to pursue this approach is made in the current exercise.[12]

4.3.3 "Immigrants"

It is not clear how to interpret their absence from the first two surveys, which contain the same number of interviewed households. One possibility is that these are individuals arrived in the villages after March 1998. However, there is no difference in the proportion of immigrants moving in treatment and control villages, neither among poor nor non-poor ones. One would expect poor households to move strategically to treatment localities, while this does not seem to occur. Moreover, sampled households do not appear to be particularly mobile, in the sense that no out-migration from either

[11]And it is even smaller when comparing means, rather than distributions.

[12]For further details on the use of RDD for the Progresa data, see Skoufias and Buddelmeyer (2003).

type of locality is recorded in the data (poor in control localities may be expected to move to treatment ones). An explanation consistent with these facts may be that this group is composed of newly-formed households. Yet, immigrants and 1997 poor have similar number of children per household. Alternatively, these households may have not been included in the first two rounds of the survey because of administrative errors. Whatever their origin, this latter group will be labelled "immigrants" in the current analysis.

No household belonging to this group and living in Progresa villages receives the treatment. Still, it may be interested to test whether there are significant differences in variables directly affected by Progresa (school enrolment, consumption, health among others) among households in control and treatment communities. The estimation of the program impact for this group is straightforward, provided that their arrival in Progresa and control localities (or their absence from the first two waves) happens on a random basis.

$$E[y|PI = t] = E[y|PI = t, C = f]$$

In this way one is estimating the overall general equilibrium impact of the program on "immigrants" including potential differences between those in control and treatment communities. However, one wouldn't know whether the impact is due to ex-ante underlying differences between control and treatment "immigrants", or to "true" general equilibrium effects. Testing for absence of underlying differences between treatment and control "immigrants" is hindered by the absence of pre-program observations. Nevertheless, the fact that there is no difference in the proportion of "immigrants" in treatment and control villages is consistent with the hypothesis of random omission of this group from the first two rounds of interviews. The differences may be in quality, rather than size of "immigrant" families in the various communities. Informative predetermined household characteristics (such as member age, gender and household size) may be used to compare first poor "immigrants" in treatment and control localities. Table 4.3 shows means of some demographic variables and reports the p-values of the test of mean equality between immigrants in control and Progresa villages.

Table 4.3: Means and p-values of differences, 1998 data

	hh size	hhh age	hhh sex[a]	age [6;16]	sons [6;16]	sons [0;5]	max grade	hectar pp
Control	4.76	39.49	0.78***	10.41	1.33	0.96	3.32	0.17
s.e.	2.31	16.73	0.41	0.12	1.61	0.90	2.41	0.55
Treatment	4.54	40.31	0.71***	10.54	1.18	0.91	3.53	0.32
s.e.	2.29	17.99	0.45	0.10	1.55	0.93	2.55	2.90
p-value	0.111	0.444	0.008	0.397	0.391	0.111	0.117	0.284

[a]: proportion males. School grade and mean age do not differ also for males and females separately.

Comparison of mean household characteristics of control and treatment poor "immigrants" reveals that there is no significant difference in demographics, apart from the fact that there are 7 percentage points more female-headed families in control villages. Quantity of land owned has a large variation among households in treatment communities and does not differ significantly between the two groups.

The lack of differences in the above variables provides some support for the hypothesis that "immigrants" in control and treatment localities do not differ in observable characteristics unrelated to the program. Using this fact, and the absence of differences in the proportion of such households in control and treatment villages one may suppose that these households do not differ in unobservables either. Since the evidence in favour of this hypothesis is quite weak, I will condition on a set of household observable characteristics in the empirical application, as explained below .

4.3.4 Non-compliers

Addressing the issue of non-compliers requires the introduction of either identifying assumptions or exclusion restrictions, unless either non-compliance occurs on a random basis, or it is possible to predict with probability one which households among the control group are the appropriate counterfactual.

No explanation is provided for households' non-compliance. Looking at pre-program patterns of school enrolment and child labour does not help. Indeed, schooling enrolment rates do not differ substantially between non-compliers and beneficiaries, as shown by the table below. Attendance rates are lower among non-compliers, though 90 percent of enroled children have a sufficiently high attendance rate to be entitled to the grant. Moreover, there are no apparent differences in schoolchildren's extra-curricular working activities. Nearly twice as many *densificado* non-compliers report being hit by some shock during the year, compared to beneficiaries, while among "1997" poor, non-compliers are hit by shocks less often than beneficiaries. Skoufias *et al.* (1999a) suggests that true non-compliers are households who were about to leave by the end of 1998, although quite a few of them are still in the sample in November 1999 (round 5). This is partially confirmed by the comparison of the 1998 and 1999 November surveys, from which it may be noticed the higher "emigration" (or disappearance) rates among non-compliers.

It is possible that non-compliance may be linked to parental absence from the household or ill health. In 1997, fathers from non-compliers' households are 13 percent more likely to be away than treated beneficiaries. However, mothers are as likely to be present. In November 1998, instead, no non-complying household has absent parents. There is a higher illness rate, although this refers only to the 4 weeks prior the interview.

Non-compliers are unevenly distributed across states. 70 percent of the 360 "1997" non-complying households live in the state of Guerrero (a quarter the size of "1997" beneficiaries living in the same state) and 18 percent in Veracruz. Guerrero and Veracruz are the states with the highest poverty level, as measured by the marginalization index.[13] *Densificado* non-compliers are

[13]With an average level of 1.31 and 0.72, while the wealthiest state, Michoacan, averages

Table 4.4: Enrolment, attendance, child labour and shock rates for non-compliers and beneficiaries

	1997		*Densificado*	
Variables (%)	beneficiary	non-complier	beneficiary	non-complier
enrolment	86	81	83	87
low attendance	2	3	9	9
working schoolchildren	1	1	4	3
shock[a]	44	41	33	70
Sample size				
September 1997	4682	360	2491	48
November 1998	4682	360	2491	48
November 1999	4188	208	2272	25

[a] : idiosyncratic shock at the household level. Percentages are from the September 1997 wave.

more evenly distributed.

In any case, no set of variables that predicts the likelihood of not participating to the program with probability one could be found. Thus, unless one is willing to ignore non-compliers altogether, some assumptions are needed to estimate the TT effect for 1997 poor. The abundance of estimable parameters suggests several candidates, as shown below.

$$E[y|P97 = 1, C = n] - E[y|P97 = 0, C = n] = 0 \tag{4.9}$$

$$E[y|NP = 1] - E[y|NP = 0] \tag{4.10}$$

$$E[y|PI = 1] - E[y|PI = 0] \tag{4.11}$$

One may assume that them effect of Progresa on non-compliers is null, as in Bloom (1984). This corresponds to (4.9) above. Alternatively, they may be the same as those for non-poor, as in (4.10). This as suggested

0.07.

by Albarran and Attanasio (2001). However, given the different wealth of this group, a better approximation may be obtained estimating the effect of the program on immigrants, as in (4.11). As already mentioned, the main limitation of this approach is that there is no pre-program information for this group. One concluding remark: the use of forgotten densificados does not seem appropriate, because the households in this group were informed of their inclusion in the eligible set. Thus, it is likely that their behaviour, especially in the initial stages of the program implementation, may have changed to comply with the program requirements.

4.3.5 Economic relevance of the estimable parameters

The possibility to separately estimate the effect of the program on the three groups has interesting economic and policy relevance. In fact, first of all it permits to estimate TT parameters for the whole sample, as well as for its sub-groups individually. Second, we can distinguish the direct program impact on different types of poor that vary both in the degree of poverty and in the program component they most benefit from: "1997" households are poorer and have more eligible children than *densificados*, so they may comparatively benefit more from the school subsidy than the latter group. Table 4.5 shows the group means of some demographic and wealth-related variables, as well as the overall poverty measure computed by the Progresa staff to separate poor from non poor households.

Indeed, *densificado* households (both forgotten and actually treated ones) have a different demographic composition. from 1997 poor. Their household head is on average ten years older; subsequently, they have fewer young children, especially of pre-school age. Furthermore, 1997 and *densificado* poor differ also in their overall poverty level. The latter group has a higher income (computed from earnings from main occupation), land and number of rooms than the original group of poor. All these variables are calculated in per capita terms. They are also consistent with differences in the poverty

Table 4.5: Household characteristics by group

Group	hh size	hhh age	hhh sex	sons [6,16]	sons [0,6)	income pp	hectares pp	room pp	Poverty index
Poor 97	5.91	42.55	0.92	1.85	1.04	2027.27	0.32	0.31	639.13
	(2.44)	(15.02)	(0.28)	(1.70)	(1.05)	(2317.54)	(0.71)	(0.27)	(82.54)
Beneficiary	4.36	52.60	0.83	1.01	0.38	2966.42	0.66	0.60	788.24
Densificados	(2.69)	(18.51)	(0.38)	(1.45)	(0.70)	(3675.69)	(1.89)	(0.64)	(72.85)
Forgotten	4.00	52.25	0.83	0.74	0.31	4272.06	0.63	0.70	834.50
Index<-.093	(2.37)	(17.71)	(0.38)	(1.20)	(0.62)	(5875.97)	(1.67)	(0.63)	(92.76)
Forgotten	4.50	52.32	0.86	0.99	0.27	3213.32	0.75	0.59	810.83
$[-.093; 1.29]$	(2.68)	(16.05)	(0.34)	(1.35)	(0.60)	(3677.84)	(1.29)	(0.50)	(70.70)

Note: hh=household; hhh=household head; pp=per person. Standard errors in parentheses. 1997 data.

indicator [14] levels computed by the Progresa staff (lower values indicate a higher degree of indigence). Hence, a comparison of TT effects for *densificados* and 1997 poor is a first step towards testing for the heterogeneity of program impact[15].

Lastly, the possibility to estimate average impacts for various types of non-participants may provide a better understanding of the impact of development programs on the receiving communities. An example of the types of applications for these parameters is the effect of Progresa on consumption smoothing. Progresa may help households smooth consumption through several channels: by increasing households' income; by providing a certain stream of future earnings (at least for a limited number of years) that may be used as collateral to borrow; by increasing liquidity at the locality level, hence the possibility to lend money. However, by looking at the effect on beneficiaries only, one cannot disentangle the individual impact of the above

[14] For details on the computation of the overall poverty index, see Skoufias et al. 1999a and 1999b.

[15] Although given that these two groups differ along at least two dimensions,what may be driving the differences would not be clear.

effects. Instead, by looking at the relationship between Progresa and consumption smoothing for the various sample sub-groups, additional information may be obtained. The effect on immigrants would measure the role played by the liquidity injection in the village. The one on forgotten households would sum the former to the effect of the entitlement to the program as a means to borrow. The estimation of these parameters may shed light on important issues for governments of developing countries which have to decide how to redistribute scarce resources efficiently.

4.4 An application: the effect of Progresa on school enrolment

In the remaining part of the paper, I estimate $TTEs$ of Progresa on child school enrolment. Progresa provides scholarships to attend grades 3 to 9, corresponding to the last four years of primary school ad the three years of junior high school. The subsidies increase with school grade and are higher for females than males at the secondary school level in order to particularly encourage female school attendance. Transfers are made on a bi-monthly basis, provided that the eligible recipient has attended at least 85 percent of classes, as confirmed by the teacher. It has been estimated that the size of the largest school subsidy corresponds to 44 percent of the male day-labourer's wage and approximately two thirds of the child full-time wage (Schultz, 2004).

The effect of Progresa on school enrolment is analyzed by Schultz (2004), who compares enrolment rates of 1997 poor in treatment and control villages, providing estimates of the "intention to treat" effect, rather than $TTEs$. The current analysis complements Schultz's by estimating the program effect on all groups of households discussed above and presenting both intention to treat and TT effects under alternative hypotheses. The possibility to estimate program effects for different types of actual beneficiaries, for eligible households who did not receive the subsidies and for non eligible families living in treatment communities provides interesting insights on the heterogeneity of treatment and general equilibrium effects. *Densificados* differ from 1997 poor along two dimensions: they have fewer children and a lower poverty level. Hence, they are likely to have a comparatively lower advantage from complying with the Progresa eligibility requirements. The effect on children from forgotten households is ambiguous, and it largely depends on whether families expect to begin receiving the missing grants soon or not. Enrolment is going to change according to one's perception or knowledge of whether and when the late payments are going to occur. Children in non-eligible poor families (immigrants) do not benefit directly from Progresa. Any significant difference may be attributable to indirect program effects. They may work in either direction (higher enrolment may decrease resources per pupil, hence teaching quality, although the extra funds to schools in Progresa areas may offset this phenomenon).

Progresa subsidies are expected to increase school enrolment of eligible children. However, it is likely to observe an enrolment increase also for the earlier primary school grades, as there are (weaker) financial incentives also for school attendance of younger pupils. Indeed, a child starting primary school in 1998 would reach the subsidized grade by 2000 (families are told that the program may only last for three years).

The program effect is expected to be larger for secondary than for primary school, because of the lower enrolment rate in the former, and for females than males, because of the larger school grants for girls. Given these considerations, I performed the analysis both for pooled children and separate genders. I also group children by their age[16], creating three different categories: children aged 6 to 9, 10 to 13, and 14 to 16. Moreover, I only consider children who were living in the household at the interview time.[17] I further restrict the sample to children who are sons or daughters

[16]Grouping children by age rather than by grade completion is preferred, because the latter is a choice variable that is influenced by the program availability. Comparing children by school level would result in different individuals being compared over time.

[17]I drop all those children who were reported to be away or to live in the household

(natural and adopted), nephews or nieces or grandchildren of the household head, excluding more distant relations and married individuals.[18] I use the September 1997, October/November 1998 and November 1999 waves (rounds 1,3 and 5) in order to have one observation per child for the three different academic years, of which the first one is prior to the beginning of the program.

4.4.1 Mean difference in enrolment rates for all sample subgroups

Pre-program enrolment rates vary considerably between age group. Virtually all children aged 6 to 9 are enrolled in 1997, while a little less than 90 percent among children aged 10 to 13 are. The lowest enrolment rates are for the eldest individuals, varying between 40 and 50 percent. Enrolment means are reported in the Appendix. The gender gap, virtually non-existent for the youngest group, varies proportionally with age, it approaches the 10 percentage point for children aged 14 to 16. Note that my classification of children excludes married individuals, hence it is likely that female enrolment rate is lower if one includes wives as well. 1997 rates do not seem to differ substantially among the various household groups, including non-poor.

Table 4.6 shows differences in average enrollment rates between treatment ad control groups before and during the program implementation by age group and gender.

only temporarily.

[18]I consider the relationship to the household head because a sizeable proportion of individuals in the valid age group is either married or has children.

Table 4.6: Difference in enrollment rates by age group, year and gender

		1997			1998			1999		
		All	Male	Female	All	Male	Female	All	Male	Female
All poor in	Diff [6;9]	-0.0014	-0.0037	0.0005	0.0036	0.0045	0.0028	**0.0052**	0.0052	0.0052
Progresa		[0.0024]	[0.0027]	[0.0035]	[0.0036]	[0.0049]	[0.0034]	**[0.0030]***	[0.0032]	[0.0041]
villages	n	15387	7820	7562	15487	7904	7904	14616	7437	7158
	Diff [10;13]	0.001	0.0124	-0.012	**0.0349**	**0.0396**	**0.0294**	**0.0335**	**0.0307**	**0.0354**
		[0.0107]	[0.0124]	[0.0136]	**[0.0108]*****	**[0.0127]*****	**[0.0125]****	**[0.0096]*****	**[0.0112]*****	**[0.0112]*****
	n	15558	8015	7540	14876	7603	7273	13766	7048	6691
	Diff [14;16]	0.0259	0.0255	0.0221	**0.0708**	**0.0766**	**0.0632**	**0.0604**	**0.0519**	**0.0721**
		[0.0244]	[0.0291]	[0.0276]	**[0.0239]*****	**[0.0285]*****	**[0.0267]****	**[0.0244]****	**[0.0295]***	**[0.0253]*****
	n	10953	5887	5064	10151	5368	4783	9248	4919	4298
Poor 97	Diff [6;9]	-0.0036	-0.0051	-0.0026	0.0064	0.0062	0.0065	0.006	0.0053	0.0067
		[0.0029]	[0.0033]	[0.0038]	[0.0046]	[0.0055]	[0.0045]	[0.0039]	[0.0041]	[0.0054]
	n	10930	5572	5355	10731	5482	5249	10267	5216	5046
	Diff [10;13]	0.0068	0.0104	0.0021	**0.0447**	**0.0469**	**0.042**	**0.0507**	**0.0474**	**0.0531**
		[0.0120]	[0.0135]	[0.0168]	**[0.0118]*****	**[0.0144]*****	**[0.0144]*****	**[0.0108]*****	**[0.0124]*****	**[0.0130]*****
	n	10008	5193	4813	9554	4916	4638	9220	4710	4501
	Diff [14;16]	0.0063	0.0067	0.0037	**0.0874**	**0.0797**	**0.0926**	**0.0835**	**0.0612**	**0.1148**
		[0.0279]	[0.0333]	[0.0322]	**[0.0276]*****	**[0.0325]****	**[0.0315]*****	**[0.0279]*****	**[0.0334]***	**[0.0309]*****
	n	5528	2996	2530	5591	3001	2590	5569	2987	2563

table continued on next page

table continued from last page

		1997			1998			1999		
		All	Male	Female	All	Male	Female	All	Male	Female
Beneficiary	Diff [6;9]	0.0076	-0.0087	0.0245	**-0.0092**	-0.0119	-0.0065	0.0053	0.0062	0.0043
		[0.0109]	[0.0068]	[0.0202]	**[0.0052]***	[0.0092]	[0.0046]	[0.0105]	[0.0122]	[0.0161]
	n	1219	612	607	1148	582	566	1063	538	522
densificados	Diff [10;13]	-0.0382	-0.0386	-0.0425	-0.0023	-0.0036	-0.0002	0.0085	0.0211	-0.0063
		[0.0239]	[0.0261]	[0.0378]	[0.0208]	[0.0264]	[0.0282]	[0.0192]	[0.0226]	[0.0280]
	n	1515	770	745	1289	674	615	1206	645	557
	Diff [14;16]	-0.0686	-0.0175	-0.1299	-0.0049	-0.0322	0.0187	-0.0249	-0.0631	0.0239
		[0.0598]	[0.0712]	[0.0792]	[0.0528]	[0.0656]	[0.0695]	[0.0539]	[0.0665]	[0.0662]
	n	1175	640	535	1068	571	497	883	464	419
Forgotten	Diff [6;9]	**0.0179**	0	**0.0352**	-0.0062	-0.0157	0	**0.0183**	**0.0137**	**0.0231**
$[-.093; 1.29]$		**[0.0106]***	[0.0000]	**[0.0204]***	[0.0063]	[0.0144]	[0.0000]	**[0.0080]****	**[0.0078]***	**[0.0123]***
	n	757	385	372	714	343	371	630	297	332
	Diff [10;13]	0.0165	0.0197	0.0013	**0.0628**	**0.0668**	0.0536	0.0045	0.0342	-0.0308
		[0.0394]	[0.0347]	[0.0695]	**[0.0202]*****	**[0.0222]*****	[0.0352]	[0.0327]	[0.0276]	[0.0574]
	n	955	499	456	814	450	364	759	408	348
	Diff [14;16]	0.107	**0.1644**	0.0233	0.1219	**0.1473**	0.107	**0.1208**	**0.1695**	0.0712
		[0.0770]	**[0.0990]***	[0.0985]	[0.0763]	**[0.0871]***	[0.0981]	**[0.0722]***	**[0.0832]****	[0.0943]
	n	781	444	337	703	350	353	595	320	275

table continued on next page

table continued from last page

		1997			1998			1999		
		All	Male	Female	All	Male	Female	All	Male	Female
Forgotten	Diff [6;9]	0.0052	0.0128	-0.003	**-0.0173**	-0.0055	**-0.03**	-0.0022	-0.0044	0.0001
Index< −.093		[0.0062]	[0.0097]	[0.0081]	**[0.0080]****	[0.0085]	**[0.0120]****	[0.0041]	[0.0045]	[0.0070]
	n	887	474	413	865	439	426	847	434	412
	Diff [10;13]	0.0247	0.0585	-0.0135	-0.0121	0.0055	-0.0325	**-0.0755**	-0.0707	**-0.0795**
		[0.0383]	[0.0539]	[0.0397]	[0.0310]	[0.0458]	[0.0282]	**[0.0363]****	[0.0571]	**[0.0343]****
	n	1004	525	478	908	458	450	778	401	377
	Diff [14;16]	-0.0429	0.0051	-0.0937	0.0149	0.0626	-0.0453	0.0581	0.0115	0.0973
		[0.0520]	[0.0718]	[0.0660]	[0.0530]	[0.0687]	[0.0707]	[0.0525]	[0.0694]	[0.0706]
	n	1081	532	549	803	446	357	578	301	275
Immigrants	Diff [6;9]				-0.0037	-0.0016	-0.0061	0.0024	0.0082	-0.0063
					[0.0088]	[0.0139]	[0.0060]	[0.0077]	[0.0124]	[0.0063]
	n				607	320	287	536	287	249
	Diff [10;13]				0.0005	-0.0079	0.0086	0.027	0.0211	0.0249
					[0.0272]	[0.0348]	[0.0379]	[0.0296]	[0.0377]	[0.0445]
	n				508	262	246	460	228	231
	Diff [14;16]				-0.0088	-0.0736	0.0533	0.0286	-0.0013	0.0719
					[0.0720]	[0.0917]	[0.0949]	[0.0778]	[0.1035]	[0.1041]
	n				272	137	135	256	138	117

table continued on next page

table continued from last page

		1997			1998			1999		
		All	Male	Female	All	Male	Female	All	Male	Female
Non poor	Diff [6;9]	0.0076	-0.0026	0.0171	0.0043	0.0097	-0.001	0.0071	0.0098	0.0044
		[0.0100]	[0.0073]	[0.0178]	[0.0121]	[0.0231]	[0.0040]	[0.0046]	[0.0061]	[0.0059]
	n	2050	994	1054	1786	915	871	1612	820	790
	Diff [10;13]	-0.0091	0.0136	-0.0316	0.0132	**0.0456**	-0.0193	0.0299	-0.0115	0.0669
		[0.0190]	[0.0245]	[0.0250]	[0.0198]	**[0.0269]***	[0.0262]	[0.0280]	[0.0337]	[0.0434]
	n	2627	1290	1337	2233	1094	1139	1821	894	921
	Diff [14;16]	**0.0644**	0.0309	**0.1006**	**0.0816**	**0.1411**	0.0245	0.0505	**0.0897**	0.0017
		[0.0386]*	[0.0524]	**[0.0488]****	**[0.0383]****	**[0.0470]*****	[0.0520]	[0.0392]	**[0.0511]***	[0.0494]
	n	2855	1570	1285	2102	1030	1072	1675	874	798

Standard errors clustered at the village level. *,**,*** significant at 10, 5, 1%.

Pre-program enrolment rates for poor households in treatment and control communities do not differ, with the exception of forgotten poor in the most marginalized communities, who have higher enrolment rates in treatment than control villages. The same occurs among non-poor children in the oldest age group. Although the differences in enrolment rates for *densificados* and forgotten households in the least marginalized villages are never significant, their point estimates are quite large.

I repeated the computation of mean differences conditioning on a set of observable characteristics: both size of the point estimates and standard errors are lower, although the differences remain not statistically different from zero. The pre-program differences for forgotten household from the most marginalized communities are no longer statistically significant for the youngest age group, when controlling for observable sources of heterogeneity. Conditioning on observable characteristics does not change the value of the point estimates and the significance levels for non-poor households, instead. One additional fact worth noticing is the absence of significant differences in immigrants' enrolment rates.

4.4.2 Double-differenced effect of Progresa on school enrolment by sample sub-group

Given the pre-program enrolment differences and the fact that some groups of households are present only in subsets of the initial village sample[19], I present difference in difference probit estimates of the marginal effect of Progresa on child enrolment.

$$\begin{aligned} P(s_{it} = 1) &= P(s^*_{it} > 0) \\ s^*_{it} &= \alpha + \beta y_t + \gamma P_{it} + \delta_{98} y_{98} P_{it} + \delta_{99} y_{99} P_{it} + \theta X_{it} + u_{it} \end{aligned} \tag{4.12}$$

where P_{it} indicates whether the child belongs to the treatment ($P_{it} = 1$) or control group ($P_{it} = 0$); y is 0 for the pre-program year, and 1 otherwise; X are a set of conditioning variables that may capture any source

[19]Hence reducing the strength of the randomisation assumption.

of difference in observable characteristics and improve the precision of the estimates.

The δ coefficients are the parameter of interest, measuring the effect of Progresa on school enrolment of treated children in 1998 and 1999. (4.12) is estimated separately by age group and gender for each of the five groups.

The additional conditioning variables are: primary school quality, measured by the ratio of resident children per teacher; number of secondary schools (standard and televised) in the locality (note that most villages do not have a secondary school); poverty level[20]; household size; household head gender, education and presence in the household. These variables are added to all specifications. Given the larger sample size for 1997 and pooled poor, I added also and spouse education and presence in the household, interacted by gender, and presence of disabled individuals in the household.

Not all individuals can be matched in the three waves. Part of it is pure sample attrition caused by individuals entering and exiting the valid age interval. A further, endogenous cause is domestic or international migration. Indeed, the previous chapter shows that labour migration, especially international, is different between treatment and control groups for 1997 poor, although the volume of migration is quite small. I estimated (4.12) using alternatively the pooled sample of all individuals aged 6 to 16, and the subset of children present in all three waves. The differences in the values of the parameters of interest between these two groups are negligible.[21]

Table 4.7 presents estimates of double-difference program effects for all groups of households in the Progresa sample. The first panel provides estimates of the average effect of the program on all poor in Progresa community (including non-compliers, forgotten households and immigrants). It is a very broad measure of the program effect, and it is the one more directly comparable to Schultz's, with whose results it shares the overall message that the impact of the program is larger for older children (secondary schoolchildren

[20]As measured by the principal component estimated by the Progresa staff with 1997 data.

[21]Results not presented but available upon request.

according to Schultz's groupings). The larger effect for "old" females than males is apparent in 1999 only.

The second and third panels provide separate estimates of the intention to treat effect for 1997 poor and treated *densificados*[22]. Such effects appear to vary substantially in terms of their broad policy implication: whereas 1997 poor seem to benefit considerably from Progresa, with enrolment going up by 3 to 7 percentage point for the whole sample, reaching a peak of 12 percentage points for girls aged 14 to 16, there is no significant effect for *densificado* children, with the exception of a one percentage point increase in enrolment of the youngest group in 1999. I repeated the estimations pooling 1998 and 1999 data, i.e. creating a dummy that takes the value of one for the two post-1997 observations, and adding additional sets of conditioning variables to try and increase the precision of the estimates, but they were never significantly different from zero. The large difference in program effect by group is further investigated below.

The absence of a significant effect for forgotten household is consistent with the hypothesis that, by November 1998, this group had realized they were not going to receive the schooling subsidy, and behaved accordingly. Somewhat less intuitive is the estimated average program effect for non-poor in Progresa villages. Enrolment for primary school males significantly increases by one percentage point for boys aged 6 to 9, and by 6 percentage points for females aged 10 to 13, while it drops dramatically, by 14 percentage point, for older girls. A possible explanation consistent with these figures may be that the larger funding to primary school offsets the disincentive caused by the higher school attendance, making primary schooling more attractive to all children. This may be possible because of the small increase in primary school enrolment, which was already close to 100% in the pre-program year. The surge in school attendance of 14 to 16 years old girls may have had the opposite effect, not only because of the decrease in

[22]Intention to treat and TT are virtually identical for *densificados* because of the low proportion of non-compliers

school quality per pupil, but also due to an increase in the relative demand for girl-specific jobs that the higher enrolment of indigent girls has caused.

Table 4.7: Double-difference program effect on enrollment rates by age group, year and gender

		[6;9]			[10;13]			[14;16]		
		All	Male	Female	All	Male	Female	All	Male	Female
All poor in	DD98	0.0037	**0.006**	0.001	**0.0205**	**0.0243**	0.0157	**0.0477**	**0.0621**	0.0341
Progresa		[0.0027]	**[0.0029]****	[0.0029]	**[0.0088]****	**[0.0098]****	[0.0124]	**[0.0198]****	**[0.0272]****	[0.0280]
Villages	DD99	**0.0047**	**0.0068**	0.0022	**0.0292**	**0.0245**	**0.0336**	**0.0369**	0.0158	**0.063**
		[0.0021]**	**[0.0022]*****	[0.0026]	**[0.0081]*****	**[0.0097]****	**[0.0113]*****	**[0.0203]***	[0.0276]	**[0.0288]****
	n	42843	21709	20868	41529	21234	20239	28678	15339	13322
Poor 1997	DD98	**0.0064**	**0.0089**	0.0028	**0.0317**	**0.0302**	**0.0337**	**0.0712**	**0.0789**	**0.0681**
		[0.0028]**	**[0.0028]*****	[0.0034]	**[0.0091]*****	**[0.0101]*****	**[0.0128]*****	**[0.0241]*****	**[0.0314]****	**[0.0334]****
	DD99	**0.0045**	**0.0071**	0.0007	**0.0423**	**0.0392**	**0.0451**	**0.0716**	0.0341	**0.1222**
		[0.0025]*	**[0.0024]*****	[0.0038]	**[0.0089]*****	**[0.0100]*****	**[0.0130]*****	**[0.0233]*****	[0.0332]	**[0.0325]*****
	n	30814	15685	14564	27891	14318	13546	16208	8766	7429
Treated	DD98	-0.0023	0.0024	-0.0026	0.0051	0.0246	-0.0188	0.0394	0.0579	0.0154
Densificados		[0.0131]	[0.0149]	[0.0121]	[0.0236]	[0.0236]	[0.0361]	[0.0546]	[0.0778]	[0.0762]
	DD99	**0.0092**	0.0099	0.0049	0.0089	0.0235	-0.0028	0.0626	0.0609	0.0857
		[0.0053]*	[0.0082]	[0.0122]	[0.0228]	[0.0225]	[0.0396]	[0.0601]	[0.0799]	[0.0831]
	n	3319	1529	1651	3793	1969	1820	3041	1657	1384

table continued on next page

table continued from last page

		[6;9]			[10;13]			[14;16]		
		All	Male	Female	All	Male	Female	All	Male	Female
Forgotten	DD98	-0.0093	-0.0176	-0.0048	0.0074	0.0295	0.009	0.0029	0.0286	-0.01
$[-.093; 1.29]$		[0.0067]	[0.0144]	[0.0050]	[0.0416]	[0.0399]	[0.0743]	[0.0881]	[0.1093]	[0.1120]
	DD99	**0.0193**	0.0239	0.0131	-0.0446	-0.0087	-0.0594	0.1033	0.0973	0.1164
		[0.0106]*	[0.0159]	[0.0104]	[0.0430]	[0.0411]	[0.0829]	[0.0850]	[0.0995]	[0.1191]
	n	2017	989	1027	2387	1272	1112	1989	1059	930
Forgotten	DD98	-0.0417	-0.0121	-0.022	-0.0059	-0.0191	-0.0095	0.0792	0.0933	0.0656
Index<-.093		[0.0367]	[0.0129]	[0.0144]	[0.0387]	[0.0495]	[0.0465]	[0.0584]	[0.0907]	[0.0740]
	DD99	-0.0047	-0.0205	0.0005	-0.0517	-0.0564	-0.0725	0.034	0.0132	0.0614
		[0.0127]	[0.0131]	[0.0012]	[0.0591]	[0.0803]	[0.0620]	[0.0642]	[0.0978]	[0.0928]
	n	2401	1255	1185	2586	1338	1247	2305	1211	1092
Non poor	DD98	0.0003	-0.0002	0.0009	0.0097	0.0186	0.0018	-0.0014	0.0938	-0.097
		[0.0056]	[0.0053]	[0.0039]	[0.0202]	[0.0251]	[0.0300]	[0.0431]	[0.0593]	[0.0596]
	DD99	**0.0071**	**0.0104**	0.0012	0.0199	-0.0374	**0.0601**	**-0.079**	-0.0221	**-0.1427**
		[0.0028]**	**[0.0031]*****	[0.0034]	[0.0192]	[0.0355]	**[0.0217]*****	**[0.0476]***	[0.0647]	**[0.0641]****
	n	5286	2621	2663	6378	3168	3208	6440	3394	3044

Robust standard errors in brackets. *,**,*** significant at 10, 5, 1%.

Table 4.8: TTEs of Progresa on school enrolment of 1997 poor

	[6;9]			[10;13]			[14;16]		
	all	male	female	all	male	female	all	male	female
Bloom									
98	0.009	0.014	0.005	**0.034**	0.030	**0.038**	**0.068**	**0.076**	**0.064**
	[0.007]	[0.009]	[0.006]	**[0.014]**	[0.017]	**[0.020]**	**[0.022]**	**[0.031]**	**[0.030]**
99	0.007	**0.012**	0.002	**0.043**	**0.040**	**0.047**	**0.069**	0.039	**0.116**
	[0.006]	**[0.006]**	[0.007]	**[0.013]**	**[0.016]**	**[0.015]**	**[0.021]**	[0.031]	**[0.028]**
Using effect on non-poor - assumption (4.10)									
98	**0.009**	**0.014**	0.004	**0.034**	**0.030**	0.038	**0.065**	**0.075**	**0.062**
	[0.004]	**[0.006]**	[0.005]	**[0.015]**	**[0.014]**	[0.022]	**[0.022]**	**[0.029]**	**[0.029]**
99	0.007	**0.013**	0.002	**0.043**	**0.040**	**0.047**	**0.069**	0.037	**0.113**
	[0.004]	**[0.006]**	[0.005]	**[0.013]**	**[0.015]**	**[0.020]**	**[0.020]**	[0.028]	**[0.032]**
Using effect on immigrants - assumption (4.11)									
98	0.009	0.014	0.005	**0.034**	**0.030**	**0.038**	**0.069**	**0.078**	0.063
	[0.007]	[0.009]	[0.004]	**[0.007]**	**[0.012]**	**[0.016]**	**[0.028]**	**[0.025]**	[0.045]
99	0.007	**0.013**	0.002	**0.043**	**0.040**	**0.047**	**0.070**	0.040	**0.114**
	[0.006]	**[0.007]**	[0.007]	**[0.011]**	**[0.006]**	**[0.018]**	**[0.025]**	[0.023]	**[0.038]**

Effects and standard errors computed by block bootstrap, block = village. Estimates in bold are significant at 5%.

Table 4.8 shows estimates of average TT effects for 1997 poor and *densificados* under alternative assumptions. The point estimates for the three sets of the $TTEs$ are virtually identical and very similar to the intention to treat estimates, because of the small proportion of non-compliers. The results are consistent with previous findings from this literature: the effect of Progresa differs by age group and gender. It is higher for older cohorts, certainly because primary school pre-program enrolment is extremely high, hence there is very limited scope for improvements; probably, also because of the positive relationship between the size of the scholarships and the school grade pupils enrol to. Attendance of higher schooling grade is associated with larger cash transfers (although the higher monetary incentives for older children are paired with a higher market value of their labour, if productivity increases with age).

The same applies to female versus male effects: Progresa has no significant effect on the enrolment of the youngest females (while it has a positive, though small effect for boys). The effect for girls aged 14 to 16 is instead a few percentage point larger than for males of the same age, in 1999. The comparative magnitude of these effects is possibly explained by the female lower pre-program enrolment rate, and by the fact that the value of secondary school grants for girls exceeds the one for boys.

The effect of Progresa on school enrolment appears to be remarkably stable over time for children aged 6 to 13. For children aged 14 to 16, instead the magnitude of the effect decreases over time for males and increases for females. The timing of the program effect is a topic that deserves further investigation.

4.4.3 The effect of poverty on enrolment for 1997 poor and treated *densificados*

The comparison of $TTEs$ for 1997 poor and treated *densificados* shows positive and significant estimates for the former, and no significant ones for the latter. In quite a few cases the point estimates between the two groups do not differ considerably, but *densificados'* estimates have much larger standard errors. While this lower precision is probably attributable to a smaller sample size, it is nevertheless useful to test whether the different poverty levels of the two groups result into different program effects. For instance, the least poor among eligible households may have enough resources to fund the migration of some of its offspring, and be less responsive to the school subsidies.

In order to test this hypothesis, I estimate (4.12) again for these two groups of children, adding the 1997 household-specific indicator of poverty level interacted by dummies for Progresa villages and years of program implementation. Table 4.9 reports the marginal effects of the poverty index for treated children in the post-1997 years. The poverty index varies between 180 and 1380, and is such that higher values are associated lower degree of indigence.

Table 4.9: Effect of household poverty level on the likelihood of being enrolled, marginal effects

	[6;9]			[10;13]			[14;16]		
	All	Male	Female	All	Male	Female	All	Male	Female
1997	-0.001	-0.005	0.003	-0.002	**0.013**	**-0.017**	0.003	0.024	-0.011
poor	[0.002]	[0.004]	[0.003]	[0.005]	[0.007]*	[0.008]**	[0.015]	[0.019]	[0.020]
Treated	0.001	**0.014**	-0.004	**0.056**	**0.046**	**0.064**	-0.027	-0.037	-0.003
densificados	[0.004]	[0.007]**	[0.006]	[0.020]***	[0.027]*	[0.033]*	[0.034]	[0.052]	[0.049]

Poverty index multiplied by 100. Standard errors clustered at the village level.

The estimated coefficients show that this is not the case. In fact, less indigent *densificado* (those with a higher index level) are either as likely or more likely to have their children enrolled in school.

4.5 Conclusion

This note explains how to estimate program effects in the Progresa data in spite of the administrative errors that result in the exclusion of a large subgroup of the eligible sample from the program (labelled forgotten household). In fact, it suggests an approach that turns the mistake into an econometric advantage, permitting the estimation of a series of parameters with clear policy relevance.

The key to the identification of counterfactuals in the presence of "forgotten" households is considering the various classes of households that belong to the eligible poor separately. It has been shown how indigent households in treatment communities are in fact composed by families classified as poor during the first stage of selection, in 1997; by households initially designed as non-poor, but later re-classified as poor, labelled *densificados*; by households who begin to be recorded in the data set only in October-November 1998 (called immigrants). Once divided in this way, it is immediately noticeable that nearly all forgotten households belong to the latter two classes - only 7 families out of the approximately 7800 "1997" poor do not receive the treatment, although they comply with the program requirements.

I suggest to use the marginalization index at the village level to identify the potentially excluded villages in the control group. I compare the distribution of pre-program variables for forgotten households in control and treatment villages with the same marginalization level. There is a reduction in the number of statistically significant differences with respect to the pooled sample of eligible households. No immigrant is an actual beneficiary.

I discuss ways to estimate the program effect for the small fraction of non-compliers, having failed to find a set of variables that permit to identify a valid counterfactual among the control group.

I provide an empirical application, estimating the effect of Progresa on school enrolment. Simple mean comparison of pre-program enrolment rates for forgotten households presented significant differences for households in control and treated villages, in three out of 27 cases. Conditioning on observables makes the differences not significant.

I present double-differenced estimates of the program $TTEs$ for the two groups of eligible poor. The conclusions reached are broadly consistent with those of the existing literature: Progresa improves school enrolment of children of all ages, but its largest impact is for older children, most likely secondary school pupils, and it is higher for girls than boys in its second year of implementation. The effect is not statistically significant for textitdensificados, although the point estimates are similar to the ones for 1997 poor. I explore the possibility that the effect of Progresa is lower for textitdensificados than for 1997 poor because the former group is less indigent than the latter, but poverty appears either unrelated or negatively correlated with the likelihood of school enrolment.

Both groups of poor households who do not receive the program for exogenous reasons have show no significant difference between treatment and control children.

4.6 Appendix

The results of the comparison of the distribution of variables for *densificado* beneficiary and forgotten households are reported below. Standard errors are computed accounting for clustering of observations at the village level. The following tables contain most of the variables from the 1997 survey (with few derived variables, such as domestic and international migration, household income, land size, household member age composition and child schooling and enrollment levels) and the consumption module from the March 1998 pre-program survey.

Only 1997 education-related variables are used, as it is feared that the March 1998 results may partially reflect some anticipation effect. School attendance may vary before the program implementation to insure that as many children are entitled to the subsidy. This includes strategic grade progression or failure.

I use the following tests: Pearson's chi squared test of association, for categorial variables; Kolmogorov-Smirnov, for continuous variables; t-test, for dichotomous variables. P-values of the differences between treatment and control groups are reported in tables below.

The columns refer to the following subgroups: column (1) is for 1997 poor. All other columns are for*densificados.* Column (2) compares beneficiary and forgotten; column (3) beneficiaries and control groups; column (4) control and forgotten groups; columns (5) and (6) compare beneficiaries and forgotten with households in the control group who live in villages within the same marginalization index intervals; columns (7) and (8) compare forgotten and control households who live in the least poor and the poorest communities, respectively (using the marginalization index value -0.093 as threshold).

Child school enrolment variables are computed considering sons and daughters, nieces/nephews, grandchildren and great-grandchildren of household heads.

Table 4.10: Mean comparison of individual and household-level variables, 1997 data

Variable	P value							
	1-3	5-6	2-5	2-6	2-5cs	2-6cs	2-6¡-.1 ¡-.1	2-6 (-.1;1.3)
Children age distr [0-17]	0.895	0.775	0.912	0.907	0.947	0.834	0.613	0.162
For children aged [6;16]:								
School attendance	0.468	0.029	0.465	0.150	0.130	0.046	0.544	0.000
Highest school grade	0.022	0.000	0.000	0.208	0.086	0.206	0.036	0.254
Highest grade (attendance)	0.061	0.010	0.004	0.127	0.186	0.136	0.044	0.260
Highest grade (no attendance)	0.235	0.000	0.004	0.148	0.567	0.193		
Worked last week	0.000	0.000	0.000	0.235	0.000	0.242	0.077	0.233
For household heads:								
Worked last week	0.001	0.000	0.000	0.000	0.000	0.001	0.007	0.024
Work frequency	0.000	0.047	0.006	0.095	0.000	0.254	0.347	0.004
Social security	0.323	0.103	0.003	0.002	0.489	0.003	0.001	0.012
Income	0.000	0.000	0.000	0.003	0.000	0.022	0.619	0.017
Age at first job	0.191	0.112	0.147	0.112	0.356	0.143	0.133	0.100
Ever migrated	0.624	0.704	0.153	0.253	0.000	0.065	0.154	0.099
Highest school grade	0.441	0.005	0.230	0.074	0.018	0.109	0.261	0.012
Household-level data:								
p40a1 DIF grant	0.683	0.3896	0.083	0.586	0.276	0.081	0.000	0.031
p40b1 Children solidarity grant	0.269	0.957	0.854	0.799	0.790	0.743	0.853	0.270
p40c1 INI grant	0.491	0.706	0.268	0.495	0.097	0.349	0.868	0.209
p40d1 PROBECAT grant	0.846	0.833	0.455	0.309	0.920	0.325	0.306	0.323
p40e1 PET	0.056	0.647	0.231	0.241	0.013	0.077	0.387	0.159
p40f1 DIF breakfast	0.362	0.387	0.226	0.072	0.200	0.000	0.000	0.283
p41a1 free tortilla	0.871	0.318	0.048	0.881	0.060	0.921	0.477	0.314
p41b1 LICONSA grant	0.884	0.977	0.816	0.861	0.953	0.680	0.453	0.000
p42a any blind	0.316	0.687	0.985	0.608	0.598	0.544	0.277	0.786
p42b any mute	0.366	0.631	0.227	0.399	0.580	0.195	0.271	0.283
p42c any deaf	0.997	0.238	0.460	0.618	0.506	0.722	0.474	0.569
p42d anyone with mental problem	0.366	0.931	0.302	0.311	0.116	0.299	0.567	0.876
p42e anyone with missing limbs	0.736	0.398	0.135	0.560	0.169	0.560	0.408	0.712
p42f anyone handicapped	0.723	0.374	0.561	0.729	0.275	0.856	0.213	0.149

Table 4.11: Mean comparison of household-level variables, 1997 data

Variable	P value							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	p97	bf	cb	cf	bcs	fcs	fl	fr
p53 house type	0.000	0.647	0.244	0.315	0.498	0.260	0.219	0.141
p54 house tenancy	0.001	0.001	0.358	0.003	0.707	0.005	0.135	0.001
p55 Floor material	0.000	0.000	0.000	0.000	0.006	0.000	0.026	0.000
p56 Roof material	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
p57 Walls material	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.000
p58 house room number	0.432	0.000	0.001	0.007	0.792	0.044	0.138	0.001
p59 water piping in land plot	0.061	0.062	0.352	0.005	0.000	0.000	0.000	0.000
p60 water piping in house	0.893	0.567	0.912	0.386	0.281	0.346	0.103	0.071
p61 bathroom in the house	0.444	0.106	0.153	0.864	0.000	0.942	0.362	0.000
p62 bathroom with running water	0.887	0.390	0.173	0.534	0.911	0.400	0.289	0.038
p63 electric light	0.590	0.000	0.000	0.543	0.000	0.102	0.014	0.147
p64 electricity meter	0.838	0.800	0.509	0.634	0.357	0.413	0.345	0.901
p65a blender	0.100	0.000	0.000	0.497	0.000	0.970	0.001	0.350
p65b refrigerator	0.671	0.000	0.000	0.616	0.011	0.730	0.103	0.676
p65c gas heater	0.707	0.000	0.000	0.143	0.003	0.000	0.000	0.392
p65d gas water boiler	0.649	0.003	0.002	0.764	0.539	0.644	0.506	0.524
p65e radio	0.806	0.000	0.000	0.753	0.000	0.516	0.009	0.040
p65f CD player	0.883	0.167	0.227	0.843	0.493	0.860	0.389	0.549
p65g television	0.188	0.000	0.000	0.699	0.000	0.821	0.000	0.003
p65h VCR	0.285	0.019	0.082	0.647	0.384	0.699	0.256	0.833
p65i washing machine	0.581	0.003	0.007	0.516	0.001	0.223	0.094	0.498
p65j electric fan	0.175	0.321	0.167	0.761	0.026	0.199	0.233	0.461
p65k car	0.385	0.097	0.920	0.070	0.376	0.011	0.245	0.081
p65l van	0.941	0.032	0.981	0.009	0.498	0.000	0.000	0.001
p66 land ownership	0.327	0.566	0.328	0.739	0.013	0.262	0.282	0.000
p67 number of plots	0.021	0.825	0.683	1.000	1.000	1.000	0.984	1.000
land size	0.000	0.346	0.016	0.264	0.025	0.310	0.007	0.491
p71 animal ownership	0.244	0.598	0.979	0.599	0.607	0.188	0.000	0.000
p72a horses	0.379	0.069	0.196	0.450	0.844	0.301	0.249	0.571
p72b donkeys	0.789	0.004	0.243	0.079	0.370	0.025	0.702	0.000
p72c oxen	0.572	0.703	0.308	0.199	0.269	0.230	0.795	0.137
p73a lambs and sheep	0.111	0.997	0.493	0.418	0.352	0.432	0.612	0.954
p73b cows and veal	0.327	0.192	0.863	0.140	0.323	0.132	0.078	0.678
p73c hens	0.906	0.778	0.115	0.213	0.583	0.197	0.012	0.388
p73d pigs	0.744	0.562	0.066	0.208	0.030	0.255	0.021	0.098
p73e rabbits	0.669	0.480	0.351	0.026	0.299	0.041	0.150	0.182

Household-level variables, continued

Variable	P value							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	p97	bf	cb	cf	bcs	fcs	fl	fr
Last week, how many days did you eat:								
p034a01 tomato	0.000	0.021	0.000	0.176	0.014	0.111	0.053	0.548
p034a02 onion	0.964	0.139	0.005	0.126	0.023	0.117	0.018	0.264
p034a03 potato	0.692	0.004	0.001	0.462	0.994	0.310	0.254	0.026
p034a04 carrot	1.000	0.000	0.000	0.876	1.000	0.999	0.909	1.000
p034a05 green vegetables	0.886	0.792	0.104	0.844	0.995	0.508	0.142	0.056
p034a06 orange	0.010	0.488	0.669	0.584	0.000	0.671	0.465	0.000
p034a07 banana	0.012	0.780	0.072	0.001	0.017	0.001	0.057	0.000
p034a08 apple	0.995	0.004	0.005	1.000	0.922	1.000	0.967	1.000
p034a09 lemon	0.745	0.908	0.731	0.218	1.000	0.139	0.102	0.028
p034a10 other vegetables	0.385	0.008	0.062	0.851	0.015	0.851	0.774	0.140
p034b01 corn tortilla	1.000	1.000	0.998	1.000	0.969	1.000	0.855	0.999
p034b02 nixtamalque bread	0.982	0.994	0.620	0.999	0.995	0.952	0.625	0.999
p034b03 corn	0.578	1.000	0.996	0.833	0.188	0.775	0.641	0.049
p034b04 white bread	1.000	0.039	0.209	0.509	0.953	0.744	1.000	0.004
p034b05 sweet bread	0.989	0.058	0.955	0.000	0.001	0.000	0.088	0.184
p034b06 sliced bread	0.908	1.000	0.991	1.000	1.000	1.000	1.000	0.993
p034b07 wheat flour	0.721	0.813	0.988	0.994	0.921	0.993	0.964	0.162
p034b08 soup	0.614	0.364	0.754	0.671	0.890	0.790	1.000	0.989
p034b09 rice	0.976	0.114	0.022	0.912	0.146	0.861	0.050	0.637
p034b10 crackers	0.040	0.002	0.867	0.020	0.034	0.049	0.969	0.317
p034b11 beans	0.000	0.583	0.000	0.000	0.073	0.000	0.000	0.002
p034b12 cereal	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
p034c01 chicken	0.015	0.647	0.013	0.346	0.000	0.333	0.654	1.000

Consumption and expenditure, continued

Variable	P value							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	p97	bf	cb	cf	bcs	fcs	fl	fr
p034c02 beef/pork	0.830	0.697	0.314	0.999	0.103	1.000	1.000	0.983
p034c03 lamb/goat	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
p034c04 fish	0.004	1.000	0.986	0.922	0.294	0.890	1.000	1.000
p034c05 tuna	0.974	1.000	1.000	0.848	0.841	0.842	1.000	1.000
p034c06 egg	0.699	0.015	1.000	0.000	0.838	0.000	0.000	0.756
p034c07 milk	0.183	0.000	0.004	0.264	0.565	0.489	0.861	0.037
p034c08 butter pork	0.238	1.000	1.000	1.000	1.000	1.000	0.280	0.094
p034d01 pastelitos	1.000	1.000	1.000	0.999	1.000	0.997	1.000	0.919
p034d02 soda	0.281	1.000	0.842	0.493	0.523	0.492	0.379	1.000
p034d03 alcoholic drinks	1.000	0.999	1.000	0.986	1.000	0.834	0.834	1.000
p034d04 coffee or tea	0.842	0.179	0.045	0.769	0.000	0.911	0.915	0.957
p034d05 sugar	0.977	0.946	0.563	0.999	0.689	0.981	1.000	0.754
p034d06 vegetable oil	0.250	0.912	0.667	1.000	0.362	1.000	0.554	0.557
How did you get the								
p035a01 chili	0.460	0.512	0.046	0.274	0.002	0.267	0.704	0.002
p035a02 onion	0.482	0.002	0.020	0.122	0.001	0.160	0.949	0.005
p035a03 potato	0.049	0.089	0.025	0.002	0.001	0.001	0.424	0.006
p035a04 carrot	0.462	0.618	0.477	0.588	0.347	0.569	0.542	0.604
p035a05 green vegetables	0.434	0.328	0.468	0.344	0.531	0.330	0.268	0.638
p035a06 orange	0.564	0.092	0.508	0.184	0.000	0.269	0.259	0.000
p035a07 banana	0.021	0.000	0.295	0.000	0.000	0.000	0.766	0.096
p035a08 apple	0.774	0.905	0.790	0.495	0.827	0.484	0.187	0.461
p035a09 lemon	0.000	0.000	0.421	0.000	0.001	0.002	0.681	0.080
p035a10 other vegetables	0.156	0.744	0.009	0.002	0.000	0.001	0.095	0.195
p035b01 corn tortilla	0.429	0.004	0.660	0.002	0.002	0.004	0.774	0.256
p035b02 nixtamalque bread	0.242	0.277	0.198	0.249	0.017	0.292	0.160	0.951
p035b03 corn	0.694	0.176	0.871	0.216	0.052	0.198	0.786	0.384
p035b04 white bread	0.499	0.924	0.222	0.125	0.235	0.146	0.217	0.040
p035b05 sweet bread	0.322	0.784	0.001	0.005	0.003	0.007	0.072	0.140
p035b06 sliced bread	0.992	0.290	0.035	0.350	0.247	0.298	0.430	—
p035b07 wheat flour	0.677	0.177	0.315	0.401	0.453	0.385	0.326	0.694
p035b08 soup	0.812	0.044	0.768	0.150	0.731	0.128	0.195	0.452
p035b09 rice	0.973	0.478	0.451	0.179	0.586	0.190	0.288	0.554
p035b10 crackers	0.071	0.297	0.513	0.366	0.509	0.260	0.238	0.161
p035b11 beans	0.045	0.575	0.031	0.073	0.003	0.293	0.122	0.380
p035b12 cereal	0.380	0.867	0.840	0.999	0.911	1.000	0.480	0.789

Consumption and expenditure, continued

Variable	P value							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	p97	bf	cb	cf	bcs	fcs	fl	fr
p035c01 chicken	0.131	0.085	0.174	0.121	0.992	0.098	0.042	0.254
p035c02 beef/pork	0.968	0.855	0.292	0.356	0.358	0.345	0.162	0.283
p035c03 lamb/goat	0.268	0.098	0.008	0.497	0.065	0.512	0.547	0.523
p035c04 fish	0.570	0.003	0.146	0.222	0.499	0.167	0.724	0.387
p035c05 tuna	0.343	0.368	0.474	0.757	0.806	0.710	0.361	0.187
p035c06 egg	0.869	0.001	0.022	0.259	0.188	0.231	0.574	0.163
p035c07 milk	0.240	0.082	0.607	0.114	0.864	0.115	0.257	0.139
p035c08 butter pork	0.082	0.654	0.883	0.402	0.465	0.444	0.239	0.710
p035d01 pastelitos	0.832	0.365	0.703	0.338	0.352	0.340	0.364	—
p035d02 soda	0.238	0.411	0.534	0.657	0.194	0.700	0.977	0.464
p035d03 alcoholic drinks	0.432	0.290	0.188	0.652	0.399	0.634	0.001	0.040
p035d04 coffee or tea	0.426	0.000	0.000	0.289	0.017	0.328	0.899	0.423
p035d05 sugar	0.152	0.051	0.056	0.199	0.039	0.182	0.588	0.090
p035d06 vegetable oil	0.348	0.154	0.349	0.125	0.061	0.113	0.348	0.083
p036a01 chili	0.953	0.301	0.015	0.016	0.222	0.191	0.007	0.358
p036a02 onion	0.380	0.207	0.035	0.361	0.710	0.373	0.007	0.494
p036a03 potato	0.647	0.282	0.041	0.272	0.082	0.236	0.222	0.367
p036a04 carrot	0.197	0.794	0.655	0.681	0.583	0.688	0.550	0.783
p036a05 green vegetables	0.128	0.087	0.387	0.235	0.383	0.244	0.132	0.727
p036a06 orange	0.130	0.542	0.282	0.959	0.536	0.982	0.112	0.516
p036a07 banana	0.194	0.673	0.420	0.858	0.931	0.871	0.375	0.925
p036a08 apple	0.834	0.896	0.248	0.276	0.612	0.191	0.065	0.761
p036a09 lemon	0.867	0.315	0.068	0.552	0.378	0.569	0.305	0.218
p036a10 other vegetables	0.046	0.070	0.993	0.041	0.865	0.029	0.123	0.252
p036b01 corn tortilla	0.335	0.108	0.032	0.440	0.299	0.419	0.051	0.693
p036b02 nixtamalque bread	0.573	0.320	0.055	0.799	0.509	0.704	0.019	0.223
p036b03 corn	0.581	0.279	0.097	0.545	0.474	0.576	0.038	0.859
p036b04 white bread	0.024	0.759	0.859	0.948	0.485	0.921	0.662	0.766
p036b05 sweet bread	0.094	0.432	0.774	0.464	0.671	0.435	0.043	0.841
p036b06 sliced bread	0.396	0.932	0.136	0.116	0.338	0.120	0.360	0.057
p036b07 wheat flour	0.607	0.706	0.315	0.127	0.186	0.118	0.115	0.444

Consumption and expenditure, continued

Variable	P value							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	p97	bf	cb	cf	bcs	fcs	fl	fr
p036b08 soup	0.537	0.394	0.250	0.792	0.617	0.768	0.069	0.125
p036b09 rice	0.253	0.342	0.290	0.449	0.356	0.441	0.116	0.369
p036b10 crackers	0.253	0.212	0.606	0.754	0.097	0.747	0.320	0.626
p036b11 beans	0.160	0.058	0.017	0.347	0.212	0.332	0.082	0.614
p036b12 cereal	0.656	0.562	0.890	0.538	0.288	0.563	0.620	0.348
p036c01 chicken	0.055	0.067	0.297	0.076	0.458	0.082	0.249	0.077
p036c02 beef/pork	0.321	0.315	0.117	0.305	0.258	0.315	0.090	0.370
p036c03 lamb/goat	0.628	0.421	—	0.282	—	0.287	0.335	—
p036c04 fish	0.311	0.304	0.336	0.095	0.776	0.108	0.036	0.675
p036c05 tuna	0.108	0.369	0.347	0.173	0.567	0.193	0.416	0.406
p036c06 egg	0.368	0.193	0.149	0.515	0.084	0.507	0.090	0.196
p036c07 milk	0.131	0.023	0.102	0.894	0.751	0.873	0.704	0.134
p036c08 butter pork	0.497	0.610	0.011	0.008	0.063	0.010	0.029	0.214
p036d01 pastelitos	0.359	0.004	0.462	0.087	0.831	0.103	0.312	0.468
p036d02 soda	0.097	0.310	0.301	0.812	0.114	0.840	0.725	0.954
p036d03 alcoholic drinks	0.743	0.918	0.166	0.129	0.859	0.108	0.015	0.590
p036d04 coffee or tea	0.579	0.192	0.454	0.552	0.238	0.539	0.141	0.199
p036d05 sugar	0.318	0.306	0.169	0.744	0.464	0.747	0.268	0.241
p036d06 vegetable oil	0.263	0.227	0.447	0.348	0.623	0.330	0.060	0.300
Last week, why didn't you eat								
p037a05 green vegetables	0.000	0.736	0.060	0.001	0.136	0.005	0.000	0.003
p037a07 banana	0.266	0.044	0.920	0.003	0.867	0.007	0.006	0.014
p037b09 rice	0.224	0.060	0.737	0.057	0.680	0.085	0.018	0.717
p037c01 chicken	0.008	0.124	0.141	0.000	0.341	0.001	0.014	0.003
p037c06 egg	0.078	0.333	0.760	0.217	0.329	0.335	0.446	0.007
p037c07 milk	0.002	0.005	0.001	0.008	0.096	0.017	0.023	0.081
p037d02 soda	0.082	0.019	0.060	0.001	0.128	0.005	0.004	0.922

Consumption and expenditure, continued

Variable	P value							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	p97	bf	cb	cf	bcs	fcs	fl	fr
How much did you spend?								
p038a how much did you spend?	0.735	0.000	0.137	0.005	0.011	0.017	0.265	0.030
p038b how much did you spend?	0.022	0.963	0.414	0.810	0.900	0.745	0.996	0.067
p038c how much did you spend?	0.053	0.117	0.008	0.783	0.045	0.676	0.959	0.344
p038d how much did you spend?	0.560	0.058	0.086	0.643	0.044	0.706	0.395	0.531
p039a how much do you owe of this?	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
p039b how much do you owe of this?	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
p039c how much do you owe of this?	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
p039d how much do you owe of this?	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
p040 how many meals do you eat at home?	0.063	0.278	0.259	0.353	0.005	0.331	0.408	0.000
ep041 do all family members eat at home?	0.589	0.445	0.454	0.920	0.514	0.798	0.436	0.042
ep042 how many people eat outside?	0.904	0.002	0.608	0.005	0.654	0.011	0.019	0.573
Do those who eat outside								
p04301 take any food	0.201	0.030	0.557	0.270	0.692	0.242	0.357	0.246
p04302 eat at relatives' house	0.367	0.695	0.313	0.488	0.411	0.488	0.657	0.712
p04303 receive food as part of their salary	0.425	0.894	0.815	0.921	0.377	0.932	0.697	—
p04304 eat at school	0.022	0.138	0.478	0.231	0.251	0.254	0.534	—
p04305 receive food by anyone	0.234	0.454	0.349	0.167	0.377	0.173	0.376	—
p04306 buy it	0.193	0.199	0.491	0.504	0.623	0.512	0.281	0.272
p044 How much do they spend on food?	0.669	0.998	0.425	0.512	0.975	0.512	0.461	—
p045 How much can you spend per week?	0.034	0.048	0.039	0.887	0.292	0.927	0.121	0.628
p046 How much do you spend per week?	0.474	0.127	0.000	0.250	0.281	0.260	0.002	0.049

Consumption and expenditure, continued

Variable	P value							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	p97	bf	cb	cf	bcs	fcs	fl	fr
If you had more money, would you spend it in								
p04701 food	0.110	0.003	0.067	0.082	0.001	0.123	0.004	0.231
p04702 fixing house	0.251	0.028	0.482	0.227	0.445	0.142	0.232	0.305
p04703 clothes or shoes	0.002	0.268	0.212	0.412	0.151	0.384	0.925	0.042
p04704 pay debts	0.427	0.521	0.541	0.106	0.187	0.084	0.986	0.019
p04705 alcoholic drinks	0.165	0.269	0.616	0.175	0.333	0.808	0.423	0.909
p04706 outings and entertainment	0.264	0.674	0.318	0.183	0.181	0.228	0.498	0.470
p04707 medicine	0.036	0.662	0.100	0.140	0.211	0.112	0.351	0.081
p04708 school supplies	0.352	0.659	0.380	0.747	0.094	0.758	0.969	0.136
p04709 toys	0.148	0.136	0.144	0.139	0.451	0.137	0.307	0.081
p04710 would save it rather than spend it	0.660	0.523	0.999	0.398	0.493	0.352	0.707	0.305
Expenditure last week								
p04801 transport to school	1.000	1.000	0.999	1.000	0.991	1.000	1.000	1.000
p04802 transport to other places	0.093	0.770	0.006	0.081	0.302	0.060	0.227	0.012
p04803 cigarettes or tobacco	1.000	1.000	1.000	0.996	1.000	1.000	1.000	0.936
p04804 alcoholic beverages	1.000	1.000	1.000	1.000	0.927	0.999	0.790	1.000
p04805 non alcoholic beverages	0.124	0.330	0.013	0.649	0.000	0.708	0.996	0.232
Expenditure last month								
p04901 personal and home hygiene	0.001	0.000	0.001	0.004	0.008	0.017	0.230	0.425
p04902 medicine	0.657	0.473	0.626	0.990	0.906	0.904	0.909	0.138
p04903 medical appointments	0.825	0.152	0.312	0.958	0.994	0.996	0.996	0.222
Expenditure in the last six months								
p05001 pots and pans	0.064	0.671	0.904	1.000	0.202	1.000	0.159	0.138
p05002 toys for boys and girls	0.994	0.881	1.000	0.998	0.870	1.000	1.000	0.553

Consumption and expenditure, continued

Variable	P value							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	p97	bf	cb	cf	bcs	fcs	fl	fr
p05003 clothes/fabric for girls	0.051	0.013	0.098	0.965	0.219	0.986	0.951	0.811
p05004 clothes/fabric for boys	0.011	0.005	0.009	1.000	0.428	1.000	1.000	0.863
p05005 clothes/fabric for women	0.355	1.000	0.577	0.298	0.409	0.308	0.211	0.334
p05006 clothes/fabric for men	0.892	1.000	0.477	0.319	0.428	0.217	0.027	0.633
p05007 shoes for girls	0.739	0.001	0.047	0.527	0.364	0.600	0.895	0.510
p05008 shoes for boys	0.063	0.000	0.001	0.999	0.012	1.000	1.000	1.000
p05009 shoes for women	0.008	0.184	0.953	0.423	0.066	0.627	0.510	0.926
p05010 shoes for men	0.004	0.871	0.898	0.615	0.871	0.811	0.274	0.823
p05011 things for school	0.363	0.251	0.943	0.714	0.528	0.888	0.630	0.990
p05012 cooperation in school	0.024	0.335	0.814	0.718	0.992	0.700	0.824	0.762
If you had more money, would you spend more on								
p05101 food	0.247	0.085	0.058	0.762	0.003	0.700	0.470	0.259
p05102 fixing house	0.781	0.026	0.413	0.261	0.169	0.151	0.124	0.630
p05103 clothes or shoes	0.038	0.882	0.071	0.073	0.426	0.076	0.002	0.136
p05104 paying debts	0.176	0.507	0.257	0.487	0.623	0.410	0.069	0.503
p05105 buying animals	0.652	0.647	0.749	0.648	0.266	0.548	0.105	0.035
p05106 buying seeds/plants	0.001	0.399	0.959	0.106	0.894	0.135	0.025	0.466
p05107 tools	0.358	0.791	0.749	0.937	0.435	0.698	0.725	0.885
p05108 electrical appliances	0.884	0.866	0.853	0.922	0.387	0.616	0.811	0.358
p05109 alcoholic drinks	0.238	0.757	0.305	0.660	0.115	0.660	0.827	0.208
p05110 entertainment	0.382	0.518	0.536	0.493	0.699	0.500	0.420	0.854
p05111 medicine	0.041	0.133	0.125	0.659	0.009	0.465	0.035	0.010
p05112 things for school	0.364	0.274	0.888	0.142	0.443	0.102	0.643	0.030
p05113 toys	0.910	0.763	0.065	0.240	0.002	0.775	0.562	0.160
p05114 would rather save it	0.037	0.009	0.230	0.189	0.168	0.116	0.329	0.095

Table 4.12: Mean enrolment by age groups, all children

	1997			1998			1999		
	[6;9]	[10;13]	[14;16]	[6;9]	[10;13]	[14;16]	[6;9]	[10;13]	[14;16]
Poor 1997									
Control	0.991	0.864	0.441	0.987	0.881	0.450	0.982	0.884	0.483
	0.096	0.343	0.497	0.115	0.323	0.498	0.132	0.320	0.500
Treatment	0.989	0.874	0.454	0.993	0.922	0.533	0.988	0.935	0.575
	0.106	0.331	0.498	0.081	0.269	0.499	0.108	0.247	0.494
Densificados									
Control	0.986	0.889	0.485	0.998	0.902	0.522	0.971	0.916	0.524
	0.116	0.315	0.500	0.044	0.297	0.500	0.168	0.278	0.500
Treatment	0.983	0.891	0.437	0.991	0.913	0.515	0.972	0.928	0.524
	0.128	0.312	0.496	0.097	0.282	0.500	0.165	0.259	0.500
Forgotten right									
Control	0.989	0.881	0.476	0.998	0.897	0.539	0.973	0.916	0.530
	0.106	0.324	0.500	0.046	0.304	0.499	0.162	0.278	0.500
Treatment	1.000	0.927	0.581	0.992	0.971	0.653	1.000	0.947	0.656
	0.000	0.261	0.494	0.089	0.167	0.477	0.000	0.224	0.476
Forgotten left									
Control	0.989	0.882	0.470	0.997	0.900	0.420	0.992	0.934	0.446
	0.104	0.323	0.500	0.052	0.301	0.494	0.089	0.248	0.498
Treatment	0.994	0.880	0.433	0.980	0.879	0.470	0.995	0.879	0.475
	0.076	0.325	0.496	0.140	0.326	0.500	0.067	0.326	0.500
Immigrants									
Control	.	.	.	0.992	0.917	0.509	0.991	0.881	0.500
	.	.	.	0.088	0.276	0.502	0.094	0.325	0.502
Treatment	.	.	.	0.989	0.918	0.500	0.994	0.908	0.529
	.	.	.	0.107	0.276	0.502	0.080	0.290	0.501
Non poor									
Control	0.994	0.886	0.407	0.984	0.900	0.472	0.990	0.889	0.505
	0.078	0.318	0.492	0.125	0.300	0.500	0.101	0.314	0.500
Treatment	0.993	0.904	0.494	0.991	0.904	0.566	0.997	0.911	0.534
	0.084	0.295	0.500	0.097	0.294	0.496	0.057	0.284	0.499

Table 4.13: Mean enrolment by age groups, males

	1997			1998			1999		
	[6;9]	[10;13]	[14;16]	[6;9]	[10;13]	[14;16]	[6;9]	[10;13]	[14;16]
Poor 1997									
Control	0.992	0.879	0.479	0.987	0.897	0.475	0.982	0.897	0.513
	0.091	0.326	0.500	0.115	0.305	0.500	0.133	0.304	0.500
Treatment	0.987	0.892	0.489	0.995	0.936	0.561	0.989	0.945	0.578
	0.115	0.310	0.500	0.073	0.244	0.496	0.104	0.228	0.494
Densificados									
Control	0.997	0.934	0.486	0.996	0.908	0.556	0.967	0.931	0.584
	0.058	0.249	0.501	0.061	0.290	0.498	0.179	0.254	0.494
Treatment	0.987	0.930	0.442	0.988	0.933	0.547	0.957	0.950	0.573
	0.112	0.256	0.497	0.111	0.250	0.498	0.203	0.218	0.496
Forgotten right									
Control	1.000	0.930	0.479	0.996	0.902	0.569	0.972	0.925	0.589
	0.000	0.256	0.500	0.064	0.298	0.496	0.165	0.264	0.493
Treatment	1.000	0.976	0.645	0.982	0.988	0.750	1.000	0.961	0.753
	0.000	0.153	0.480	0.132	0.111	0.435	0.000	0.194	0.434
Forgotten left									
Control	0.984	0.897	0.463	0.995	0.874	0.391	1.000	0.923	0.428
	0.125	0.305	0.500	0.074	0.333	0.489	0.000	0.268	0.497
Treatment	0.996	0.914	0.491	0.989	0.874	0.508	0.996	0.886	0.494
	0.061	0.280	0.501	0.104	0.332	0.501	0.066	0.319	0.501
Immigrants									
Control	.	.	.	0.985	0.925	0.556	0.985	0.904	0.508
	.	.	.	0.121	0.265	0.502	0.121	0.296	0.504
Treatment	.	.	.	0.984	0.917	0.482	0.993	0.925	0.507
	.	.	.	0.127	0.277	0.503	0.081	0.264	0.503
Non poor									
Control	0.992	0.879	0.426	0.972	0.906	0.479	0.989	0.901	0.521
	0.087	0.327	0.495	0.165	0.293	0.500	0.107	0.300	0.500
Treatment	0.989	0.917	0.502	0.986	0.944	0.593	0.998	0.910	0.561
	0.106	0.277	0.500	0.120	0.231	0.492	0.045	0.286	0.497

Table 4.14: Mean enrolment by age groups, females

	1997			1998			1999		
	[6;9]	[10;13]	[14;16]	[6;9]	[10;13]	[14;16]	[6;9]	[10;13]	[14;16]
Poor 1997									
Control	0.990	0.850	0.399	0.987	0.866	0.422	0.983	0.872	0.448
	0.097	0.358	0.490	0.114	0.341	0.494	0.130	0.334	0.498
Treatment	0.991	0.855	0.411	0.992	0.906	0.498	0.987	0.923	0.574
	0.096	0.352	0.492	0.088	0.292	0.500	0.113	0.266	0.495
Densificados									
Control	0.976	0.846	0.484	1.000	0.896	0.490	0.975	0.898	0.455
	0.153	0.362	0.501	0.000	0.306	0.501	0.156	0.303	0.499
Treatment	0.980	0.850	0.432	0.994	0.891	0.472	0.987	0.902	0.476
	0.141	0.358	0.496	0.080	0.313	0.500	0.115	0.298	0.501
Forgotten right									
Control	0.977	0.834	0.471	1.000	0.892	0.509	0.974	0.905	0.462
	0.149	0.373	0.500	0.000	0.312	0.501	0.160	0.294	0.500
Treatment	1.000	0.866	0.489	1.000	0.950	0.560	1.000	0.929	0.551
	0.000	0.342	0.502	0.000	0.220	0.498	0.000	0.259	0.500
Forgotten left									
Control	0.994	0.863	0.476	1.000	0.927	0.453	0.985	0.945	0.471
	0.076	0.345	0.500	0.000	0.260	0.499	0.123	0.228	0.501
Treatment	0.992	0.845	0.362	0.969	0.885	0.419	0.995	0.872	0.451
	0.091	0.363	0.482	0.175	0.320	0.495	0.070	0.335	0.499
Immigrants									
Control	.	.	.	1.000	0.910	0.467	1.000	0.862	0.491
	.	.	.	0.000	0.288	0.503	0.000	0.346	0.505
Treatment	.	.	.	0.994	0.919	0.520	0.994	0.887	0.562
	.	.	.	0.078	0.275	0.503	0.080	0.318	0.500
Non poor									
Control	0.995	0.893	0.385	0.997	0.895	0.464	0.991	0.877	0.493
	0.068	0.309	0.487	0.055	0.307	0.499	0.095	0.329	0.501
Treatment	0.997	0.891	0.485	0.996	0.866	0.540	0.996	0.911	0.503
	0.056	0.312	0.500	0.063	0.341	0.499	0.067	0.284	0.501

Bibliography

1. Albarran, P. and Attanasio, O. (2001), "Do public transfers crowd out private transfers? Evidence from a randomized experiment in Mexico", mimeo, University College London.

2. Angrist, J., Imbens, G. and Rubin, D. (1996), "Identification of Causal Effects Using Instrumental Variables", *Journal of the American Statistical Association*, 91(434), 444-55.

3. Attanasio, O., Meghir, C. and Santiago, A. (2001), "Education choices in Mexico: using a structural model and a randomized experiment to evaluate Progresa", mimeo, University College London.

4. Bloom, H. (1984), "Accounting for no-shows in experimental evaluation designs", *Evaluation Review*, 8(2), 225-46.

5. Behrman, J.R. and Todd, P. (1999), "Randomness in the experimental samples of Progresa (education, health and nutrition program)", *mimeo*, International Food Policy Research Institute.

6. Borjas, G. (1987), "Self-selection and the earnings of immigrants", *American Economic Review*, 77(4), 531-53.

7. ——— (1994), "The economics of immigration", *Journal of Economic Literature*, 32, 1667-1717.

8. ——— and Bratsberg, B. (1996), "Who Leaves? The Outmigration of the Foreign-Born", *Review of Economics and Statistics*, 78(1), 165-76.

9. ———, Freeman, R. B. and Lang, K. (1991), "Undocumented Mexican-Born Workers in the United States: how many, how permanent?", in Abowd, J.M. and Freeman R. B., eds., Immigration, trade and the labor market. Chicago, IL: University of Chicago Press, 1991, 77-100.

10. Blundell, R. W. and Smith, R. J. (1989), "Estimation in a class of simultaneous limited dependent variable models", *Review of Economic Studies*, 56(1), 37-57.

11. Bratsberg, B. (1995), "Legal versus illegal U.S. immigration and source country characteristics", *Southern Economic Journal*, 61(3), 715-27.

12. Buddelmeyer, H. and Skoufias, E. (2003), "An evaluation of the performance of regression discontinuity design on PROGRESA", IZA Discussion Paper 827, Institute for the Study of Labor.

13. Carter, T.J., (1999), "Illegal immigration in an efficiency wage model", *Journal of International Econonomics*, 49(2), 385-401.

14. Chesher, A. and Lancaster, T. (1983), "The estimation of models of labour market behaviour", *Review of Economic Studies*, 50, 609-624.

15. Cornelius, W. (1989), "Impact of the 1986 US immigration law on emigration from rural Mexican sending communities", *Population and Development Review*, 15(4), 689-705.

16. ——— (2001), "Death at the border: efficacy and unintended consequences of US immigration control policy", *Population and Development Review*, 27(4), 661-85.

17. Da Vanzo, J. (1983), "Repeat Migration in the United States: Who Moves Back and Who Moves On?", *Review of Economics and Statistics*, 65(4), 552-9.

18. Davila, A., Pagan, J.A. and Grau, M.V. (1999), "Immigration reform, the INS and the distribution of interior and border enforcement resources", *Public Choice*, 99(3-4), 327-45.

19. Davila, A., Pagan, J.A. and Soydemir, G. (2001), "The short-term and long-term deterrence effects of INS border and interior enforcement on undocumented immigration", *Journal of Economic Behaviour and Organisation*, 49, 459-472.

20. Department of Justice. Immigration and Naturalization Service.

21. Djajic, S. and Milbourne, A. (1988), "A general equilibrium model of guest-worker migration", *Journal of International Economics*, 25, 335-51.

22. Donato, K.M., Durand, J. and Massey, D. (1992), "Stemming the tide? Assessing the deterrent effects of the IRCA", *Demography*, 29(2), 139-157.

23. Dubois, P., de Janvry, A. and Sadoulet, E. (2001), "Effects on school enrollment and performance of a conditional transfers program in Mexico", mimeo.

24. Dustmann, C. (1997) "Return migration, uncertainty and precautionary savings", *Journal of Development Economics*, 52, 295-316.

25. ——— (2003) "Return migration, wage differentials and the optimal migration duration", *European Economic Review*, 47(2), 353-67.

26. The Economist (2002), "A survey of migration", 2 November issue.

27. Entorf, H. (2000), "Rational migration policy should tolerate non-zero illegal migration flows: lessons from modelling the market for illegal migration", IZA Discussion Paper No. 199.

28. Espenshade, T. (1994), "Does the threat of border apprehension deter US immigration?", *Population and Development Review*, 20(4), 871-92.

29. ——— (1995), "Unauthorized immigration to the United States", *Annual Review of Sociology*, 21, 195-216.

30. ——— (1995b), "Using INS border apprehension data to measure the flow of undocumented migrants crossing the U.S.-Mexico frontier", International Migration Review, 29(2), 545-565.

31. Ethier, W.J. (1986), "Illegal immigration: the host-country problem", *American Economic Review*, 76(1), 56-71.

32. Faini, R. and Venturini, A. (1993), "Trade, aid and migrations. Some basic policy issues", *European Economic Review*, 37, 435-442.

33. Faini, R. and Venturini, A. (1994), "Migration and growth: the experience of Southern Europe", *CEPR Discussion Paper 964*.

34. Faini, R. and Venturini, A. (2001), "Home bias and migration: why is migration playing a marginal role in the globalization process?", *CHILD Working Paper* 27/2001.

35. Fair, R. (1970), "The estimation of simultaneous equation models with lagged endogenous variables and first order serially correlated errors", *Econometrica*, 38(3), 507-16.

36. Hanson, G., Robertson, R. and Spilimbergo, A. (2002), "Does border enforcement protect US workers from illegal migration?", *Review of Economics and Statistics*, 84(1), 73-92.

37. ——— and Spilimbergo, A. (1999), "Illegal immigration, border enforcement and relative wages: evidence from apprehensions at the US-Mexico border", *American Economic Review*, 89(5), 1337-1357.

38. ——— and Spilimbergo, A. (2001), "Political economy, sectoral shocks and border enforcement", *Canadian Journal of Economics*, 34(3), 612-638.

39. Heckman, J, Smith, J. and Taber, C. (1998), "Accounting for dropouts in the evaluation of social programs", *Review of Economics and Statistics*, 80(1), 1-14.

40. Hill, J.K. (1987), "Immigrant Decisions Concerning Duration of Stay and Migratory Frequency", *Journal of Development Economics*, 25(1), 221-34.

41. Inter-American Development Bank, (2003)

42. Katz, E. and R. and Stark, O. (1986), "Labour migration and risk aversion in less developed countries", *Journal of Labour Economics*, 4(1), 134-49.

43. Kennan, J. and Walker, J. R. (2003), "The effect of expected income on individual migration decisions", mimeo, University of wisconsin - Madison.

44. Kossoudji, S. (1992), "Playing cat and mouse at the Mexican-American border", *Demography*, 29(2), 159-180.

45. Massey, D. (1987), "The Ethnosurvey in Theory and Practice", *International Migration Review* 21, 1498-1522.

46. Massey, D. and Espinosa, C. (1997), "What's driving Mexico-US migration? A theoretical, empirical and policy analysis", *American Journal of Sociology*, 102, 939-99.

47. Massey, D. and Singer, A. (1995), "New estimates of undocumented Mexican migration and the probability of apprehension", *Demography*, 32(2), 203-13.

48. Massey, D. and Zenteno, R. (1999) "A validation of the ethnosurveys: The case of Mexico-U.S. migration", *International Migration Review*, 33, 766-793.

49. Meghir, C. and Whitehouse, E. (1997), "Labour market transitions and retirement of men in the UK", *Journal of Econometrics*, 79, 327-354.

50. Mexican Migration Project (2004), http://mmp.opr.princeton.edu/.

51. Nickell, S. (1979), "Estimating the probability of leaving unemployment", *Econometrica*, 47(5), 1249-1266.

52. Pozo, S., ed. (1986), Essays on legal and illegal immigration: Papers presented in a seminar series conducted by the Department of Economics at Western Michigan University, W.E. Upjohn Institute for Employment Research: Kalamazoo, Michigan.

53. Schick, F and Schick, R. editors (1991), Statistical Handbook on U.S. Hispanics, Phoenix, Arizona: Oryx Press.

54. Schultz, T. P. (2004), "School subsidies for the poor: evaluating the Mexican PROGRESA poverty program", *Journal of Development Economics* (forthcoming).

55. Sjaastad, Larry A. (1962), "The Costs and Returns of Human Migration", *Journal of Political Economy,* 70, 80-93.

56. Skoufias, E., Davies, B and Behrman, J.R. (1999a), "Final report - An evaluation of the selection of beneficiary households in the education, health and nutrition program (PROGRESA) of Mexico", *mimeo*, International Food Policy Research Institute.

57. Skoufias, E., Davies, B and de la Vega, S. (1999b), "Targeting the poor in Mexico: an evaluation of the selection of households into PROGRESA ", *mimeo*, International Food Policy Research Institute.

58. Smith, R. J. and Blundell, R. W. (1986), "An exogeneity test for a simultaneous equation tobit model with an application to labour supply", *Econometrica*, 54(3), 679-86.

59. Stark, O., Helmenstein, C. and Yegorov, Y. (1997), "Migrants' savings, purchasing power parity, and optimal migration duration", *International Tax and Public Finance*, 4(3), 307-24.

60. Stark, O. and Taylor, J. E. (1991), "Migration incentives, migration types: the role of relative deprivation", *Economic Journal*, 101, 1163-78.

61. United States General Accounting Office (1991), "Border Patrol. Southwest border enforcement affected by mission expansion and budget", GAO/GGD91-72BR.

62. United States General Accounting Office (1996), "Border Patrol. Staffing and enforcement activities", GAO/GGD96-65.

63. Warren, R. and Passel, J. (1987), "A count of the uncountable: estimates of undocumented aliens counted in the 1980 United States Census", Demography, 24(3), 375-393

64. Warren, R. (1994), Estimates of the unauthorized immigrant population currently residing in the United States, by country of origin

and state of residence: October 1992. Washington, DC: Statistical
Division, Immigration and Naturalization Service.