

10-22-2009

Intergenerational Transmission of Abilities and Self Selection of Mexican Immigrants

Vincenzo Caponi
Ryerson University

Follow this and additional works at: <http://digitalcommons.ryerson.ca/economics>



Part of the [Economics Commons](#)

Recommended Citation

Caponi, Vincenzo, "Intergenerational Transmission of Abilities and Self Selection of Mexican Immigrants" (2009). *Economics Publications and Research*. Paper 18.
<http://digitalcommons.ryerson.ca/economics/18>

This Working Paper is brought to you for free and open access by the Economics at Digital Commons @ Ryerson. It has been accepted for inclusion in Economics Publications and Research by an authorized administrator of Digital Commons @ Ryerson. For more information, please contact bcameron@ryerson.ca.

Intergenerational Transmission of Abilities and Self Selection of Mexican Immigrants[†]

Vincenzo Caponi[‡]

Ryerson University, Rimini Center for Economic Analysis and IZA

October 22, 2009

Abstract

This paper presents an intergenerational self selection model of migration and education that is capable of explaining the evolution of earnings and education across three generations of immigrants. By structurally estimating the model it is possible to quantify the human capital level of Mexicans in light of the self-sacrifice made by the first generation of Mexican immigrants. The results suggest that there is a significant one time loss of human capital faced by immigrants upon migration that is not transmitted to their children. Also parents with larger amounts of human capital tend to migrate more and tend to choose to remain high school educated. However, given the better educational opportunities offered in the US, they migrate with the expectation of their children becoming college educated. Therefore, measures that rely on the earnings performance and educational attainment of immigrants underestimate the amount of human capital they bring into the host country.

JEL Classification: F22, J24, J61,

[†]I am particularly grateful to Audra Bowlus and Lance Lochner, for all their support and useful suggestions. I also thank Adalbert Mayer for providing me with advice and help to develop this paper. Participants at seminars at Rochester, particularly Ronni Pavan, McMaster and at the Society for Economic Dynamics Annual Meeting in Istanbul also provided useful comments and suggestions. The referee process has been very helpful and I thank three anonymous referees for this.

[‡]Department of Economics, Faculty of Arts, Ryerson University, Toronto, Ontario, Canada. E-mail: vcaponi@ryerson.ca.

1 Introduction

This paper is motivated by observations regarding the performance of immigrants and their descendants. Solely focusing on Mexican immigrants in the US the following three facts can be observed.¹ First, there is a significant gain from migration in terms of earnings. The earnings of Mexican immigrants in the US are higher than the earnings of non-migrant Mexicans. Earnings continue to increase from the first to the second generation of Mexicans in the US, but then stop increasing or regress from the second to the third generation.² Second, conditional on education, the returns from migration are higher for high school educated than for college educated immigrants.³ Earnings increase for both groups from the first to the second generation, but more for college than for high school educated. However, from the second to the third generation, high school earnings stabilize while college earnings decrease. Third, measuring educational attainment in terms of the share of college educated individuals, Mexican immigrants are less educated than non-migrant Mexicans. Second generation Mexicans substantially improve their education compared to their parents, while attainment for the third generation is slightly lower than for the second.

An explanation of this evidence is provided, which is based on a structurally estimated intergenerational self selection model of migration and education. The results of the estimation indicate that it is important to distinguish between two functions of human capital. On the one hand human capital is used to generate earnings, on the other hand it is also transferred to future generations determining their future earnings. The estimated model shows that the first function is greatly and negatively affected by migration while the second is not. That is, immigrants face a significant loss of capacity to translate their human capital into

¹The data used are from the Current Population Survey (CPS) - March supplement. The survey collects information on foreign born individuals independently of their status of residence in the US and therefore includes information on legal as well as on illegal immigrants. As with census data, illegal immigrants in the CPS might have a lower response rate than legal ones. However, Chiquiar and Hanson (2005) show that accounting for the possible under count of illegal immigrants in the census 2000 data does not change significantly the statistics of immigrants in general.

²The definition of first generation immigrants includes all Mexican born individuals who live in the US. The definition of second generation includes all US born individuals with at least one Mexican born parent. The definition of third generation includes all individuals who identify themselves as Mexicans and are not included in the first or second generation. The third generation includes all generations of Mexicans beyond the second.

³The definition I use of high school education includes all the individuals with a high school degree or lower education. College educated are those with a high school degree that attended, for at least one term, a college.

earnings, while there is no loss of capacity to transfer human capital to their children. This distinction explains the large gap in earnings between first and second generations. Immigrants have lower earnings because they are unable to use part of their human capital to generate earnings. However, the amount of human capital transferred to the second generation is not affected by the same loss and is consequently greater than the earnings of the first generation would suggest. Moreover, second generation Mexicans do not face the same difficulty the first generation faces, they are capable of fully utilizing the inherited human capital significantly improving their performance compared to their parents. The model also indicates that the loss of human capital is higher for college educated immigrants than for immigrants with high school education or less, which explains their lower returns from migration.

This paper also suggests that Mexican immigrants are positively selected with respect to human capital. This contrasts with Borjas (1993), who, by looking at the average increase of earnings from the first to the second generation, concludes that recent cohorts of immigrants are drawn from the lower tail of the human capital distribution in their home country. However, the hypothesis proposed by Borjas that immigrants are negatively selected is inconsistent with the observations on the second and the third generation of Mexicans. Given the intergenerational persistence in the transmission mechanism of human capital, if immigrants are drawn from the lower tail of the human capital distribution, then the second generation should, on average, have higher human capital than the first, and the third higher than the second. On the contrary, this paper suggests that because of the positive self selection average human capital within the second generation is higher than the stationary average, and so are the earnings, while the third generation and successive show a reversion to the mean.

Even among the high school educated, immigrants are positively selected. Given the difficulty to adapt their own human capital to the US labor market, particularly for college educated immigrants, it may be optimal for them to avoid facing the cost of acquiring a college education in Mexico. Some of these individuals would have chosen to attend a college were they not able to migrate to the US. However, because of altruism toward future generations, migrating is an optimal choice for them due to the benefits of their children and grandchildren, who are expected to become college educated taking full advantage of the inherited human capital. This explains why, despite the positive selection, educational attainment among immigrants is

lower than among non-migrants.

Finally, the paper evaluates alternative policies that aim at integrating immigrants into the US labor market that may have positive impacts on the human capital distribution in the host country, and policies that aim at contrasting the brain drain from Mexico to the US. I propose two policies that integrate immigrants faster, one that targets high school educated and the other college educated immigrants. The simulation results under these alternative policies suggest that both policies can be effective at increasing the human capital of immigrants. However, while keeping the overall cost of the policy the same, the second policy generates larger effects on human capital than the first. By reducing the loss of human capital faced by college educated immigrants, the policy generates direct positive effects on the overall human capital and earnings of the first generation, but also indirect effects on the first and the second generation by strengthening the self selection process of immigrants. I also propose one policy that contrasts the brain drain by reducing the cost of education in Mexico. The simulation results indicate that this policy is effective in contrasting the brain drain from Mexico to the US. Immigrants become less self-selected with respect to human capital and among non-migrant Mexicans the average human capital increases. Moreover, the policy also reduces the migration rate especially among the highly educated Mexicans.

The paper is structured as follows. Section 2 briefly looks at some of the related literature. Section 3 highlights a few features from the data. Sections 4 and 5 introduce the model that is estimated and discuss its identification. Section 6 discusses the estimation results and section 7 evaluates the effects of alternative policies. Section 8 concludes.

2 Literature Review

The evidence reported in this paper relative to Mexican generations of immigrants in the US is consistent with several papers in the migration literature. Chiswick (1978) first noticed that second generation immigrants tend to perform better in terms of earnings than the first and the third generations. More recently, Card (2005) found similar evidence despite using very different data and type of immigration into the United States. Chiswick analyzes data from immigrants that migrated mainly in the period before 1965. Immigrants from this period

were subject to a quota system law passed in 1925 that favored North European immigration.⁴ Card’s analysis targets immigrants that migrated after 1965. Card reports that even among the least educated, and in particular Mexicans, second generation immigrants on average overcome 80% of the disadvantages that their parents experienced. Similar findings are also reported in the sociological literature by Glick and White (2004) and Kao and Tienda (1995). Using the National Education Longitudinal Study (NELS), they find that the second generation of immigrants is more likely to continue education to post secondary education than the first and third generation. In addition, McKeever and Klineberg (1999) report that among the Hispanics in Houston the second generation of immigrants is better educated and also has higher earnings than the first and third generation.

One contribution of my paper is to the literature on the selection mechanism of immigrants in a destination country. Borjas (1987) and Borjas (2000) found that recent cohorts of immigrants in the US are negatively selected. Among them are Mexicans, which represent about 30% of the whole foreign born population in the US. Chiquiar and Hanson (2005), on the contrary, suggest that Mexican immigrants are positively or intermediate selected. A strong positive selection with respect to human capital is also consistent with the evidence produced by a large body of research in the sociology literature. Glick and White (2004) estimate a multinomial logistic model that predicts the likelihood that a student remains with less than high school education, graduates from high school, or attends college. They find that, among immigrants in general, the second generation is more likely to go to college than the third generation, and, although the results are less robust, this is true also among Mexican immigrants.⁵ Moreover, consistently with the altruistic motive for migration and positive self selection on human capital, they find that parents’ expectations about their children’s education as well as parents’ involvement greatly increases the likelihood of enrolling in college. Similar conclusions are drawn also by Kao and Tienda (1995) who claim that foreign-born youth are at a disadvantage due to their limited English skills, but since their parents encourage them to acquire a college education, second generation youth are likely to perform well in school. However, they

⁴“The law was influenced by the research of Brigham (1923) who classified immigrants into four categories: “Nordic”, “Alpine”, “Mediterranean” and “Asian”, and argued that members of the Alpine and Mediterranean races had lower intelligence than Nordics[...]” Card (2005) footnote 1, p. 2.

⁵The level of significance of the relevant parameters is much higher for Asian descent than for Mexicans. However, this is in part due to the fact that the sample of Mexicans is less than one third of that of Asians.

find strong support that this is the case for Asian immigrants, but no support for Hispanics.

3 Sample Characteristics

This section focuses on the dynamics of earnings and education across successive generations of Mexicans in the US and compares the performance of these generations with Mexican non-migrants - those who live in Mexico. Table 1 provides a descriptive overview of earnings and educational attainment for different generations of Mexicans in Mexico and in the US. The table presents average log hourly earnings, educational attainment⁶, as well as age, years of schooling, the total population size and the number of observations. The sample used for the analysis is based on the 2000 Mexican census, for observations on individuals resident in Mexico, and on the Current Population Survey (CPS), March supplement from 1994 to 2008, pooled together.⁷

First, there is a large difference between the log earnings of Mexicans working in Mexico and Mexicans in the US. Table 1 shows that the first generation of male immigrants earns on average about 1.5 log-points more than those who remained in Mexico. There is also a large difference between first and successive generations of Mexicans in the US, with the second and third generation of males earnings about 0.37 log-points more. Table 1 also shows that the first generation has, on average, a much lower level of education than their descendants. While about 13.4% of immigrants have some college education, the percentage is more than double among the second and third generations. However, it is also interesting to note that male immigrants are less educated than those who remained in Mexico - 15.4% of them have a college education - and that the second generation is slightly more educated than the third - 45.4% compared to 43.7%. As for the earnings from the second to the third generation they slightly increase by about 2%.⁸

To better understand the relationship between education and the intergenerational evolution of the earnings of immigrants, I run a regression of log hourly earnings on a set of dummy

⁶Educational attainment is measured here as the share of the referenced population with some college education.

⁷See Appendix A for details on the sample selection.

⁸Note that males of the third generation are on average older than males of the second, a fact that might explain the earnings difference.

Table 1: Sample - Descriptive Statistics

Variable	Men				Women			
	Mexico	1 st Gen	2 nd Gen	3 rd Gen	Mexico	1 st Gen	2 nd Gen	3 rd Gen
Log Wages	0.6831 (0.0024)	2.1796 (0.0046)	2.5447 (0.0082)	2.5641 (0.0053)	0.7585 (0.0038)	1.9768 (0.0064)	2.3613 (0.0083)	2.3550 (0.0054)
College Share	0.1539 (0.0011)	0.1337 (0.0029)	0.4541 (0.0069)	0.4369 (0.0045)	0.2356 (0.0020)	0.1768 (0.0051)	0.5265 (0.0075)	0.4970 (0.0049)
Age	36.5823 (0.0349)	39.8096 (0.0898)	36.7312 (0.1629)	38.0781 (0.0954)	34.9013 (0.0470)	41.7280 (0.1327)	36.7970 (0.1747)	38.5028 (0.0997)
School Years	8.3826 (0.0147)	8.4668 (0.0342)	12.2449 (0.0299)	12.3304 (0.0187)	9.8088 (0.0224)	8.9248 (0.0534)	12.5101 (0.0315)	12.5600 (0.0182)
Population	11031430	1229532	500392	1086856	5226140	441129	388572	895240
N. obs	102448	13597	5257	12155	44136	5539	4404	10596

Author's calculation based on a sample combined of observations from the Mexican Census 2000 and CPS surveys from 1994 to 2008. See Appendix A for more details on how the sample is selected. Standard errors in parenthesis.

variables for generation and education and a set of individual characteristics, including a square function of age centered at 48 interacted with the education/generation dummies, and a set of dummies for year of birth.^{9,10} I also include dummies among the controls for survey years and for geography.¹¹ Table 2 reports the returns to migration and education, as well as the evolutions of earnings across all generations.¹² The table reports results for men and women using two definitions: the standard definition and a stricter definition. The standard definition of the second generation is the one provided above which identifies the second generation as all those US born individuals with at least one parent born in Mexico. The stricter definition imposes that a person not only has at least one parent born in Mexico, but also identifies her/himself as Mexican.

⁹The introduction of the year of birth in the regression allows to net out the effect of belonging to a particular cohort on earnings. Cohort effects can be relevant in earnings regressions with data on migrants as different cohorts can be affected in their migration decisions by the particular business cycle at the age in which migration is usually a real option. The year of birth can be used for CPS data since there are 15 pooled surveys from 1994 to 2008. The same is not possible for Mexican data.

¹⁰Similar to Altonji and Doraszelski (2005), the inclusion of a polynomial function of age centered at 48, about the middle of the working life span considered in the data, is intended to capture the evolution of earnings before and after this mid point so that the regression returns an average of past, current and future earnings, and can be interpreted as a proxies for lifetime earnings.

¹¹Clearly geography cannot be considered entirely exogenous, especially in a migration model. The reason to include geographical dummies is to take into account the different cost of living, and therefore purchasing power of earnings, in different locations.

¹²the results reported in Table 2 are computed by taking the differences between the estimated coefficients reported in Table A.1. Appendix A reports the details of the regression model and the sample used. All the estimates are based on OLS regressions.

The first four rows present the returns to college for each generation. Focusing on the first column, it can be noted that the return to college is quite high for each generation of men, but is much higher for Mexican non-migrants. Among the first generation of Mexicans in the US, first row, those men with college have an average earnings 0.38 log-points larger than those with only high school or less. The gain from college increases for the second generation, 0.45 log-points, but then decreases for the third, 0.40. In row 4, an impressive 1.22 log-points quantifies the returns to college in Mexico.¹³

Rows 5 and 6 present the returns to migration calculated by looking at the differences between the log earnings of Mexicans living in Mexico and the first generation Mexicans in the US for each schooling category. Row 5 shows that immigrants with only high school in the US earn on average 1.63 log-points more than Mexicans living in Mexico, while for college educated the gain is only 0.79. That is, the returns to migration are much lower for college educated than for high school educated Mexicans. Rows 7 to 10 show the evolution of earnings from the first to the second generation and from the second to the third. From row 7 it appears that the change from the first to the second generation for immigrants without college is highly positive, showing a gain of 0.35 log-points. The gain for immigrants with college from the first to the second generation is even larger than for the high school educated at 0.42. From the second to the third generation there is instead a contraction of earnings of about 1%, not statistically significant, for immigrants without college, and a more substantial and significant contraction of about 7% for college educated.¹⁴

The last two columns of Table 2 show that the results are robust to the definition of second and third generations. Including in the second generation only those who identify themselves as Mexicans - the stricter definition - does not change the results substantially, as can be seen by looking at the returns to college for the second generation in row 2, columns 1 and 3, and the

¹³Mexico Attanasio and Binelli (2010) in Figure 10 report a wage premium for college educated males compared to high school or less educated around 90% for the year 2000. They use micro data from the ENEU, a survey on urban employment run by the INEGI that is representative at the national level. The differences between Attanasio and Binelli estimates and mine are due to two main factors: they do not control for age, cohort and geography; their data only covers urban employment and not rural.

¹⁴In interpreting these results, particularly the difference between the first and successive generations, it should be noted that the group of high school or less educated Mexican immigrants is much more heterogeneous than the equivalently defined group of the second and third generations. Evidence of this fact is reported by Chiquiar and Hanson (2005) and Caponi (2006). Tables C.1 and C.2 in Appendix C provide more evidence of this heterogeneity and a short discussion of its implications for the interpretation of the results.

Table 2: Intergenerational Evolution of Earnings				
Dependent Var.: Log Hourly Wage	Standard Definition		Stricter Definition	
	Men	Women	Men	Women
<hr/> Returns to College <hr/>				
1 st Gen. Imm.	0.3832 (0.0232)	0.3654 (0.0341)	0.4259 (0.0197)	0.3923 (0.0291)
2 nd Gen. Mexicans	0.4546 (0.0284)	0.4510 (0.0324)	0.4389 (0.0295)	0.4505 (0.0341)
3 rd Gen. Mexicans	0.3995 (0.0175)	0.4504 (0.0193)	0.3992 (0.0171)	0.4511 (0.0189)
Mexicans in Mexico	1.2230 (0.0081)	1.1440 (0.0118)	1.2232 (0.0079)	1.1440 (0.0116)
<hr/> Returns to Migration <hr/>				
Without College	1.6263 (0.0194)	1.3507 (0.0274)	1.6231 (0.0167)	1.3629 (0.0249)
With College	0.7865 (0.0284)	0.5721 (0.0401)	0.8259 (0.0245)	0.6112 (0.0354)
<hr/> Intergenerational Dynamics <hr/>				
From 1 st to 2 nd Gen. No Coll.	0.3479 (0.0212)	0.3194 (0.0270)	0.3445 (0.0210)	0.3140 (0.0267)
From 1 st to 2 nd Gen. With Coll.	0.4193 (0.0301)	0.4050 (0.0386)	0.3575 (0.0288)	0.3723 (0.0362)
From 2 nd to 3 rd Gen. No Coll.	-0.0100 (0.0226)	-0.0317 (0.0265)	-0.0040 (0.0227)	-0.0265 (0.0268)
From 2 nd to 3 rd Gen. With Coll.	-0.0652 (0.0246)	-0.0323 (0.0269)	-0.0437 (0.0256)	-0.0260 (0.0283)

evolution of the earnings in rows 7 to 10. This gives some confidence that the bias induced by relying on self-identification as Mexicans to define the third generation, criticized by Duncan and Trejo (2005), is negligible.¹⁵ The facts shown in this section motivate the next section of the paper, which introduces a theoretical model capable of replicating the features of the data.

4 Model

Presented here is a partial equilibrium intergenerational altruistic model where a person chooses the level of education and the country of residence. Following Borjas (1993), I assume that the

¹⁵In Appendix C I discuss the results in more detail and present another set of results using data from the Latino Sample 1989-1990 of the Panel Study of Income Dynamics (PSID).

choice of the country of residence is based on the earnings capacity of an individual represented by the endowment of human capital and the alternative returns to human capital across locations. Following the vast body of research that shows the existence of multiple skills, I assume that individuals are endowed with an amount of human capital composed of two abilities that can be used alternatively depending on the acquired level of education: 1) an intellectual or college ability used to generate earnings if some college education is acquired; and 2) a manual or high school ability used in the absence of college education.¹⁶ Therefore, agents not only choose their location, but the level of schooling they want to acquire, or, in other words, which of the two abilities they want to use for producing earnings.

Acquiring an education has a cost, which is different for the two levels of education. Wages, or skill prices, are also different depending on the skill used. Further, skill prices and schooling costs differ across countries. I also assume that there is unidirectional migration from Mexico to the US, as observed in the data.¹⁷ Moreover, because there is no uncertainty about the returns to migration and its psychic cost when the decision to migrate is made, there is no return migration.

Mexican born individuals can choose from the following four options: 1) high school and working in Mexico; 2) college and working in Mexico; 3) high school and working in the US; and 4) college and working in the US. The last two choices imply that the individual migrates. The choices of American born individuals are only between the two levels of education considered; in this sense their choices are identical to those in Mayer (2008).

As in Borjas (1993), I assume that migrants face some costs of migration. However, while Borjas assumes that there is a psychic or pecuniary cost that does not affect the earnings capacity of immigrants, as in Caponi (2006) I assume that there are two costs: a psychic cost; and an ability cost. The psychic cost is assumed uncorrelated with the abilities and makes immigrants homesick for the rest of their working life. The ability cost makes part of the ability endowments of immigrants unusable for producing earnings in the host country.

¹⁶See Heckman, Stixrud, and Urzua (2006) for the role of cognitive and non-cognitive skills in explaining labor market and behavioral outcomes. See also Sattinger (1993), Heckman and Sedlacek (1990), Heckman and Scheinkman (1987) and the seminal work of Willis and Rosen (1979) for more evidence on multiple skills and self selection into education.

¹⁷This assumption is made in order to simplify the exposition of the model, but it does not have any implication as the model, given the estimated parameters, would predict that only Mexicans can have an incentive to migrate to the US and not viceversa.

Following Borjas (1993) I also assume that the second and successive generations of immigrants inherit the human capital from their parents, and use the intergenerational framework in Mayer (2008) to model the intergenerational transmission of abilities from parents to children. Therefore, the endowment of an individual depends stochastically on the endowment of his or her parent (each parent has one child). Skills are transmitted following a bivariate autoregressive process, in which each skill is allowed to be transmitted at a different rate. The process is assumed to be the same for Mexican and American born individuals and their children that stay in the parental country. However, it can be different if parents and children are born in different countries in the sense that the transmission of skills can be affected by the loss of capacity to transmit human capital faced by immigrants. The choice of schooling does not affect the endowments of ability transmitted to children. This does not rule out the possibility that parents may invest in the human capital of their children, particularly at an early age.¹⁸ Parents with higher levels of abilities transfer more ability to their children. This can be due to genetics but it can also be the result of investments in human capital.¹⁹ Moreover, because the schooling choice considered is between high school and college, the assumption does not rule out that early schooling from elementary to the completion of high school can have an effect on the development of both abilities, allowing for a causal effect of parents' early education on their children's abilities. However, by the time the decision to go to college is made, the intergenerational transmission process is completed and the person making the decision knows with certainty the amount of both abilities.²⁰

Finally, parents care about their children and they maximize their welfare given by their own lifetime utility plus the discounted welfare of their children, or, in other words, the discounted utility of the whole dynasty. Given this structure, if the Mexican born parent wishes to migrate, he or she has to take into account the loss of part of her or his abilities, the psychic cost of

¹⁸In this sense the paper does not distinguish between *nature* and *nurture*, a distinction that according to Cunha and Heckman (2007) is obsolete.

¹⁹The sector specific abilities may include intelligence, work ethic, or physical attributes; all of which may be the result of investments in human capital prior to the decision of whether to attend college.

²⁰While there is an extensive literature that documents a positive relationship between parents and children's schooling, particularly for mothers, Behrman and Rosenzweig (2002) find that this relationship disappears, or even becomes negative, for mothers when appropriately controlling for unobserved components that determine parents' schooling such as ability. The relationship remains positive between fathers' and children's education, but the measured causal effect is small. These findings are consistent with the assumptions of the model that it is primarily ability endowments that are transferred from parents to children.

moving, and the gain for future generations of being born in the host country. This gain depends on the intergenerational transfer of abilities between parents and children born in different countries.

The decision process of a Mexican born agent can be divided into two steps. In the first step, the agent, conditional on living in one location, chooses the education level that maximizes his or her utility. In the second step, the agent compares the utility he or she obtains in each location plus the discounted value of the utility of future generations, and chooses to migrate or not by choosing the higher one.

I start by defining the earnings of an individual within a dynasty with a given endowment of abilities. The earnings of an individual are proportional to the level of the ability used, which depends on the chosen level of schooling, and the skill price $\pi_{a,k}$, which depends on the country of residence $a = \{mx, us\}$ and the schooling level $k = \{H, C\}$, where H stands for high school and C for college. Immigrants' earnings are also affected by the loss of human capital that reduces the capacity to use their abilities to produce earnings. This is symbolized by zI_g , where z is the proportional amount of human capital lost upon migration, and I_g be an indicator that takes value 1 if the generation g of the dynasty migrated and 0 otherwise. Therefore, let s_k be the natural logarithm of the ability levels of an individual's endowment, the log-earnings of each generation are represented by,

$$w_{a,k,g} = \pi_{a,k} + s_{k,g} - zI_g. \quad (1)$$

The abilities endowment of an individual belonging to generation $g + 1$ depends on the endowment of her parents who belong to generation g . The mechanism that governs the intergenerational transmission of abilities within a dynasty is as follows:

$$s_{k,g+1} = b_k(s_{k,g} - \xi zI_g) + u_{k,g+1}, \quad (2)$$

where ξ represents the reduction of capacity to transmit human capital which is assumed to be proportional to the loss of human capital or abilities originally faced by the immigrant parent. The parameters b_k describe the degree of intergenerational persistence in the transmission mechanism of abilities, and $u_{k,g+1}$ are error terms assumed to be normally distributed,

$$\begin{pmatrix} u_H \\ u_C \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} \middle| \begin{pmatrix} \sigma_H^2 & \sigma_{HC} \\ \sigma_{HC} & \sigma_C^2 \end{pmatrix} \right).$$

Mexican born individuals make a joint schooling-migration decision. However, it is possible to analyze the schooling decision separately from the migration choice. In fact, given that the schooling decision, conditional on the migration decision, does not affect the state of future generations, it can be analyzed without taking into account the altruistic feature of the model. Therefore, a Mexican that remains in Mexico decides to attend college if the lifetime earnings, less the cost of college, that can be obtained by being college educated are higher than the lifetime earnings obtained by being high school or lower educated. That is, if

$$w_{mx,C,g} - \tau_{mx,C} > w_{mx,H,g}, \quad (3)$$

where $\tau_{mx,C}$ is the cost of a college education, and assuming that the cost of high school is zero for any generation. The cost of a college education is assumed to be proportional to potential lifetime earnings reflecting the importance of the foregone earnings due to the fact that college educated individuals start earning a few years later than high school educated individuals. Rewriting equation (3) using equation (1) gives,

$$\pi_{mx,C} - \pi_{mx,H} - \tau_{mx,C} > s_{H,g} - s_{C,g}. \quad (4)$$

In contrast, immigrants choose to attend college if

$$\pi_{us,C} - \pi_{us,H} - \tau_{m,C} > s_{H,g} - s_{C,g}, \quad (5)$$

where τ_m is the cost of a college education for a Mexican immigrant. The second and third generation Mexicans choose to attend college if

$$\pi_{us,C} - \pi_{us,H} - \tau_{us,C} > s_{H,g+i} - s_{C,g+i}, \quad i = 1, 2, \quad (6)$$

where τ_{us} is the cost of a college education for a US born individual. Note that there are three different costs of college: the cost of acquiring a college education for a Mexican who decides to stay in Mexico; the cost for a Mexican who decides to migrate to the US; and the cost for a US born individual, either of the second or third generation. Because the cost of education reflects foregone earnings as well as direct and psychic costs associated with going to college, it is reasonable to assume that immigrants may face different costs from both non-migrants

and second and third generations of Mexicans in the US. The direct and psychic costs faced by immigrants may be close to the ones faced by non-migrants given that the education is acquired in Mexico, while foregone earnings are comparable to the ones faced by second and third generations given that they work in the US. Therefore, the combination of the two sources of costs is likely different across the three groups.

Another important feature of the theory presented here is altruism. The migration decision made by Mexican born parents affects the state of their children and their welfare since it determines their place of birth. Therefore, a Mexican born agent decides to migrate or not depending on his or her own gain from migration as well as the effects of his or her decision on the welfare of future generations.

The value of migrating for a Mexican born agent is composed of a part that describes the gain from migrating for the current generation given by their earnings w minus a psychic cost, plus the (discounted) expected value to future generations of being born in the US. Assuming log utility, and given that one period is equivalent to one generation and that there is no need to borrow or save for the future the value is given by,

$$v_m(s_{H,g}, s_{C,g}) = \max_k \left\{ w_{us,k,g} + \beta E v_{us}(s_{H,g+1}, s_{C,g+1}) \right\} - \psi, \quad (7)$$

where v_{us} is the value of being born in the US, β is the parameter that measures altruism, and ψ is an utility cost of migrating drawn at the time the migration decision is made from a normal distribution with mean μ_ψ and variance σ_ψ^2 . The value of the psychic cost is independent across generations. In the above Bellman's equation the state space of each individual is determined by his endowment (s_H, s_C) . Moreover since the choice of education of one generation does not affect the state space of the next generation it is possible to write,

$$v_m(s_{H,g}, s_{C,g}) = \max_k \{w_{us,k,g}\} + \sum_{j=1}^{\infty} \beta^j \max_k \{E w_{us,k,g+j}\} - \psi. \quad (8)$$

The value (v_{us}) of being born in the US is given by,

$$v_{us}(s_{H,g+i}, s_{C,g+i}) = \max_k \{w_{us,k,g+i}\} + \sum_{j=1}^{\infty} \beta^j \max_k \{E w_{us,k,g+i+j}\}. \quad (9)$$

An agent that decides to remain in Mexico takes into account that his or her child will be born in Mexico and will have the opportunity to migrate in the next period. Therefore, the value for an agent of not migrating is given by his or her current earnings plus the expected value of a Mexican born agent,

$$v_{mx}(s_{H,g}, s_{C,g}) = \max_k \left\{ w_{mx,k,g} + \beta E \max [v_{mx}(s_{H,g+1}, s_{C,g+1}), v_m(s_{H,g+1}, s_{C,g+1})] \right\}, \quad (10)$$

Finally, the decision to migrate or not is made in order to maximize the following

$$v(s_{H,g}, s_{C,g}) = \max \left\{ v_{mx}(s_{H,g}, s_{C,g}), v_m(s_{H,g}, s_{C,g}) \right\}. \quad (11)$$

Equation (11) simply states that, depending on his or her ability endowments, a Mexican born agent chooses to migrate or not and the level of schooling such that the best option available is obtained.

5 Estimation Procedure

Structural estimation of the model is implemented using the method of moments. Since the moments of the model that have data counterparts cannot be analytically derived, the method of moments is implemented using model simulations. McFadden (1989) provides the theoretical foundation to the Simulated Method of Moments (SMM).

As shown in the previous section, the model studied is closely related to Mayer (2008). In particular, although Mayer estimates his model using data on Americans, the process that determines the intergenerational transmission of abilities is assumed to be identical for Americans and for Mexicans. This allows the utilization of the work already done by Mayer (2008) and the ability to use his parameter estimates while concentrating on estimating the remaining parameters related to the behavior of Mexicans. In this sense the estimation performed here can be viewed as a second stage of a Two Stage Simulated Method of Moments estimation (2SSMM) which was first proposed by Newey and McFadden (1994) and Gourinchas and Parker (2002)²¹.

²¹Newey and McFadden (1994) provide the foundation of the two step GMM procedure, while Gourinchas and Parker (2002) extend this procedure to the simulated method of moments.

More formally, let $x(u_n, \chi_0)_{n=1}^N$, be a series of observed data, and $x(u_n^s, \chi)$, $n = 1, \dots, N$ and $s = 1, \dots, S$ be a set of S series of simulated data, conditional on χ . Denote $\mu(x(u_n, \chi_0))$, or simply $\mu(x_n)$, a vector of moments of the data. The SMM procedure consists in minimizing an objective function representing a measure of the distance between moments from data observations and the simulations obtained from the model that can be represented by

$$Q(\chi) = \left[\sum_{n=1}^N \left(\mu(x_n) - \frac{1}{S} \sum_{s=1}^S \mu(x(u_n^s, \chi)) \right) \right]' W_\chi^{-1} \left[\sum_{n=1}^N \left(\mu(x_n) - \frac{1}{S} \sum_{s=1}^S \mu(x(u_n^s, \chi)) \right) \right], \quad (12)$$

where W_χ^{-1} is a matrix that defines the relative weights of the moments.

In this case $\mu(x_n)$ can be partitioned into two vectors, $m(x_n)$ and $g(x_n)$, of moments. The first vector represents moments related to observations on Americans, and the second related to observations on non-migrant Mexicans as well as first, second and third generation Mexican immigrants in the US. The set of parameters can also be partitioned into two sets θ and γ , such that the set of parameters θ does not affect the moments $m(x_n)$. These parameters are the ones that only affect the behavior of Mexicans and not the behavior of Americans. Because $m(x_n)$ and $g(x_n)$ are independent moments and, most importantly, because $m(x_n)$ is independent from θ , it is possible to estimate γ independently and use the estimates in a second stage to estimate θ . The parameters Mayer estimates, taken as coming from the first stage, are

$$\gamma = [b_H, b_C, \sigma_H, \sigma_C, \rho], \quad (13)$$

while I estimate the following set of parameters

$$\theta = [\pi_{mx,C}, \pi_{us,H}, \pi_{us,C}, \tau_{m,H}, \tau_{us,C}, \tau_{mx,C}, z, \xi, \mu_\psi, \sigma_\psi^2]. \quad (14)$$

Note that the skill prices in the US are included in the set of parameters that is assumed to not affect the moments derived from American data. To estimate the model, I use only data on Mexican generations in the US and Mexico, and I do not use data on Americans. Although I do assume that the intergenerational autoregressive process is the same for Americans and Mexicans, it is not necessary to assume that Americans and Mexicans in the US face the same set of skill prices. Neighborhood effects, discrimination, and other reasons could make a Mexican with the same abilities as an American facing a different wage. Therefore, in what

follows, $\pi_{us,k}$ should be interpreted as the skill prices for Mexicans in the US.²²

There are two more parameters in the model that cannot be identified by the data used in the estimation procedure independently from other parameters: the discount or altruism factor β , and the skill price of high school ability in Mexico $\pi_{mx,H}$. Since the skill prices can only be identified up to scale, I fix the lower skill price in Mexico to be equal to zero, *i.e.* $\pi_{mx,H} = 0$. As for the discount parameter, I assign to it the value of 0.3079.²³ Assuming that a period is about 30 years long, the value reflects a discount factor of 0.9615 per year, which would generate an interest rate equal to 0.04.^{24,25}

Given the partition of the moments and parameter vectors and taking $\hat{\gamma}$ as given from Mayer's estimation, I proceed with the second stage of the 2SSMM procedure as in Gourinchas and Parker (2002), minimizing

$$Q(\theta) = \left[\sum_{n=1}^N \left(g(x_n) - \frac{1}{S} \sum_{s=1}^S g(x(u_n^s, \theta, \hat{\gamma})) \right) \right]' W_{\theta}^{-1} \left[\sum_{n=1}^N \left(g(x_n) - \frac{1}{S} \sum_{s=1}^S g(x(u_n^s, \theta, \hat{\gamma})) \right) \right]. \quad (15)$$

Importantly, the fact the Mayer's estimates can be used in the 2SSMM context, allows me to use the information on the precision of $\hat{\gamma}$, its covariance matrix, in order to obtain correct standard errors in my estimation. Let the Jacobian of the $g(x_n^s, \hat{\theta}, \hat{\gamma})$ moment functions with respect to θ be G_{θ} , and the Jacobian of the same moment functions with respect to γ be G_{γ} . Let Ω_{γ} be the covariance matrix of the γ estimates, and Ω_g the covariance matrix of the data moments. It can be proved²⁶ that a consistent estimator of the covariance matrix of θ in a 2SSMM procedure is obtained by

²²In the first stage Mayer estimates the difference between the skill prices for Americans. Footnote 39 shows that the estimated difference of the skills' prices of the skill prices for Mexicans in the US is not significantly different from the estimated difference of the skills prices for Americans. Therefore, it cannot be excluded that the skills prices are the same.

²³To perform sensitivity analysis I re-estimate the model with alternative values for beta. The results are reported in Section 6.

²⁴Notice that the model is dynastic. Agents see their children as themselves in the future and in this sense the model can be interpreted as a classical selfish model with infinitely lived agents. Therefore, the discount factor has the same interpretations as in those models, and not as in standard OLG models with different selfish generations trading their assets.

²⁵In Section 6 I conduct sensitivity analysis on the model with alternative values of β , which shows that, while it is difficult to pin down any particular number in a wide range of positive values, the estimation clearly rejects the hypothesis that altruism has no role in the migration decision.

²⁶See Laibson, Repetto, and Tobacman (2005) for a proof based on Newey and McFadden (1994) and Gourinchas and Parker (2002).

$$\Omega_\theta = Var(\theta) = (G'_\theta W_\theta^{-1} G_\theta)^{-1} G'_\theta W^{-1} [\Omega_g + \Omega_g^s + G_\gamma \Omega_\gamma G'_\gamma] W^{-1} G_\theta (G'_\theta W_\theta^{-1} G_\theta)^{-1}, \quad (16)$$

where $\Omega_g^s = \frac{1}{S} \Omega_g$ is the simulation correction. The weighting matrix I use is obtained by inverting the data moments covariance matrix, $W^{-1} = \Omega_g^{-1}$ so that it is possible to rewrite equation (16) as follows

$$Var(\theta) = (1 + \frac{1}{S})(G'_\theta \Omega_g^{-1} G_\theta)^{-1} + (G'_\theta \Omega_g^{-1} G_\theta)^{-1} G'_\theta \Omega_g^{-1} G_\gamma \Omega_\gamma G'_\gamma \Omega_g^{-1} G_\theta (G'_\theta \Omega_g^{-1} G_\theta)^{-1}. \quad (17)$$

The part that characterizes this estimator as different from the usual SMM covariance estimator is given by $G_\gamma \Omega_\gamma G'_\gamma$, which is the contribution to the covariance matrix of the uncertainty from the first step. $Var(\theta)$ increases if the covariance Ω_γ of the first step estimates increases, and also increases if G_γ , the sensitivity of the second step moments to the first step estimates, is higher.

5.1 Identification Strategy

Mayer (2008) provides a discussion of the identification of the parameters γ in the first stage estimation. He uses data from the Panel Study of Income Dynamics (PSID). The major advantage of using the PSID in his study is that it is possible to link observations of parents to observations of children. Mayer (2008) uses observations collected between 1968 and 1976 for parents, and observations collected between 1992 and 2001 for children. He collects information on schooling and earnings for each individual. He first collects earnings information for every year and then uses the individual time series to create a measure of lifetime earnings. He uses the information on lifetime earnings and schooling on each parent-child couple to calculate the moments that are used in his SMM estimation procedure. The main moments used are the averages and variances of earnings for each generation, conditional and unconditional on schooling, the averages and variances of earnings of children conditional on the schooling of their parents, the share of college educated children and parents, and the correlation between earnings and school choices of parents and children. The variation of earnings conditional on schooling choice and the intergenerational persistence in earnings identify the parameters σ_H , σ_C , b_H and b_C . The relationship between the parents earnings and educational attainment of their children identifies the parameter ρ .

Table 3: Data Moments for the SMM Estimation: Males only

Moment	Data	s.e.
Migration Rate	0.1003	0.0008
College Mex. in Mex.	0.1539	0.0010
College 1 st Gen. in US	0.1337	0.0009
College 2 st Gen. in US	0.4541	0.0013
College 3 st Gen. in US	0.4369	0.0013
Earnings HS 1 st Gen. in US	1.6263	0.0103
Earnings HS 2 st Gen. in US	1.9742	0.0193
Earnings HS 3 st Gen. in US	1.9642	0.0129
Earnings C. Mex. in Mex.	1.2230	0.0099
Earnings C. 1 st Gen. in US	2.0095	0.0238
Earnings C. 2 st Gen. in US	2.4288	0.0210
Earnings C. 3 st Gen. in US	2.3636	0.0140

See Appendix A for a detailed explanation on how the moments are calculated.

My focus here is on the identification strategy for the second stage. The moments available for the second stage estimation are summarized in Table 3. The sample used to derive the moments in the table is the same used for Table 2.²⁷ The first moment in Table 3 is the migration rate given by the share of first generation Mexican immigrants on all Mexican born individuals. The moments in the second to the fifth rows are the shares of individuals with a college education in each generation group. The remaining rows in Table 3 show information on earnings. All of the earnings moments reported in Table 3 are expressed in log hourly earnings and are averages relative to the lowest earner group represented by non-migrant Mexicans with a high school education or lower. The second column of Table 3 shows the standard error for each moment.²⁸

Once I have all of the moments I need to be certain that the model is identified. The determination of the selection mechanism, and therefore the identification of the parameters that primarily determines it, is based on the differences between the third and the second generation moments. Since all the parameters faced by the third and the second generation are the same, the differences between the second and the third generation moments are due to the changes in the ability distribution only.²⁹

²⁷See Appendix A for a detailed description of the sample used and how the moments and the covariance matrix are derived.

²⁸See Appendix A for the details on how the covariance matrix is derived.

²⁹The model assumes a stationary intergenerational environment. Increased returns to education, lower costs

This, together with the initial conditions on the ability distribution for the generation of Mexicans non-migrants that has to be stationary, determines the distribution of abilities for all generations given that the intergenerational dynamics of the ability distribution is known from the first stage. In particular, it determines the average abilities of the first generation of migrants compared to the non-migrant Mexicans, which is the selection mechanism. Therefore, the changes of the earnings moments and the educational share from the second to the third generation identify the skill prices in the US as well as the variance of the disutility shock distribution, all the parameters that mostly affect the selection mechanism, although they are not the only ones. The skill prices are identified because the averages of the abilities distribution conditional on the schooling choice are known, the variance is identified because it determines the intensity of the changes in all three moments.³⁰ Moreover, the knowledge of the distribution of abilities across generations allows to identify the parameters from the moments on earnings and educational attainment. From the earnings of the second generation is identified the parameter describing the loss of ability to transmit human capital ξ , while the cost of a college education in the US is identified by the share of college educated. From the earnings of the first generation is identified the loss of human capital z , while from the share of the college educated is identified the cost of college for the first generation. From the moments of Mexicans non-migrants are identified the cost of a college education in Mexico and the skill price for the college skill. Finally, the average of the disutility shock is identified by the migration rate.

of communication and transportation, and immigration policies have changed in the last decades, and these changes may have affected the migration decisions of different generations of migrants. However, particularly for Mexico-US migration, the reduction of migration costs implied by better communications and transportation, may have been negligible in terms of lifetime earnings in the US. As for the legal environment, the abolition of the Bracero program in 1964 likely increased the cost of migrating to the US disproportionately to people with lower skills. However, since then illegal immigration started to be an important phenomenon, and as of today most of the immigrants entering the US from Mexico are illegal immigrants. The change in the legal environment brought an important change in the status of the immigrants, but did not change the number and type much. Finally, the skill premium increased significantly in the US, as it did in Mexico at a very similar pace, keeping the relative returns in the two countries similar over time. See for example Kurokawa (2006), who explains the common trend with increased international trade between the two countries, and Binelli (2009), who documents that the “convexification” of earnings driven by skill biased technical change has been a common phenomenon for many countries.

³⁰Figures D.1 and D.2 in Appendix D show simulations with a very low variance and with a higher variance. By looking at the figure it is possible to see that the impact of a lower variance is to strengthen the selection mechanism. The simplified model presented in the same appendix gives some analytical results that help to understand how the selection works, and it shows that, under certain parameterizations, the selection mechanism is monotonic with respect to each ability - although with opposite sign.

6 Estimation Results

Table 4 reports the estimates from Mayer (2008) that are used as first stage estimates. The table shows significant and sizable parameters related to the intergenerational transmission of abilities. The intellectual ability is shown to be more persistent ($b_C = .52$) than the manual ability ($b_H = .11$).

Table 4: First Stage Parameter Estimates

Parameters	Point Estimate	s.e.
b_H	0.1118	0.0669
b_C	0.5249	0.0417
ρ	0.5154	0.1318
σ_H	0.3776	0.0135
σ_C	0.5624	0.0346

Mayer (2008), Table 2.

The standard deviation of the shocks associated with the transmission of the intellectual ability is larger than the standard deviation of the shock of the manual ability. Together with the persistence parameters, this implies that the variance of the intellectual ability in a cross section of individuals is higher than the variance of the manual ability.³¹ Proposition 3 in Mayer’s paper proves that this, together with a strong correlation between the abilities, is a sufficient condition for the probability of children’s college attendance to be a positive function of the parents’ wage when parents are also college educated. Moreover, the probability of having college educated children is also increasing in parents’ earnings when parents are not college educated. As clarified later, this point has important consequences for the migration model presented here.

Table 5 presents the estimated parameters and the standard errors obtained by the second stage of the 2SSMM procedure.³² Model 1 in the table refers to the model discussed above and for which I discussed the identification. Model 2 is a generalization of Model 1 in which the parameter related to the loss of human capital (z) is allowed to differ by education. Finally,

³¹The variance in a cross section is given by: $\bar{\sigma}_k^2 = \sigma_k^2 / (1 - b_k^2)$.

³²Standard errors in Table 5 are obtained using equation (17). To evaluate equation (17), I needed to numerically calculate the derivatives of the moment functions with respect to both sets of parameters. I also needed the covariance matrix of the data moments, and the covariance matrix of the estimates from the first stage. The last bit of information was kindly provided by Mayer.

Model 3 is a restricted version of Model 2, where the ξ is constrained to be zero.

The last row in the table shows the inverse of the goodness of fit for each model. These values are calculated as weighted sums of squares of the deviations between the simulated and data moments, where the weights are obtained using the optimal weighting matrix.^{33,34} The fit of Model 2 is substantially better than the fit of Model 1, which implies that allowing the loss of human capital to differ by education substantially improves the explanatory power.³⁵ The restrictions in Model 3 do not change the fit much from Model 2, which supports the hypothesis that there is no loss of capacity to transmit human capital from the first to the second generation of Mexicans. The rest of the quantitative analysis focuses on the results of Model 3 as it is more parsimonious in terms of parameters compared to Model 2, and it shows a much better fit compared to Model 1.

The first two rows of the table show the price set faced by Mexicans remaining in Mexico. As stated previously, the skill price for the lower educational level is normalized to zero ($\pi_{mx,H} = 0$). Therefore, all of the skill prices are relative to it and should be interpreted as the difference from the lowest skill price. The table shows that the skill price for college educated in Mexico ($\pi_{mx,C}$) is about 0.6 log points and is significant at the 5% level. This parameters indicates the returns to college in Mexico for a randomly selected person, that is, net of self selection. The cost of a college education for a Mexican remaining in Mexico (τ_{mx}) is about 1.09.³⁶ Assuming a high school or less educated non-migrant earns about 50,000³⁷ US dollars, the return to college is about 41,000 US dollars implying an average cost of attending college of about 60,000 US

³³The optimal weighting matrix is define as $W_{opt} = [\Omega_g + \Omega_g^s + G_\gamma \Omega_\gamma G_\gamma']^{-1}$.

³⁴A formal over-identifying restriction test can be done on Models 1 and 3 by testing the null hypothesis $\eta(\hat{\theta}, \hat{\gamma}) = 0$; in both cases the model is rejected. A possible reason for rejecting is that the model is not flexible enough to capture the differences between the two types of occupations, especially among immigrants. Model 2 has as many moments as parameters and therefore cannot be tested with the same test.

³⁵In Model 1, the simulated earnings of high school educated immigrants are significantly lower than their data counterpart, while the earnings of college educated are higher. In Models 2 and 3, these differences vanish almost completely. This is shown in Table B.1, Appendix B.

³⁶Table C.1 in Appendix C shows that non migrant Mexicans have the highest average years of education among college educated while they also have the lowest average years of education among the high school educated. Part of the high returns to education may be explained by heterogeneity within the groups in terms of years of schooling. In this sense, the estimated skill price for college might be upward-biased together with the cost of education.

³⁷The figure is obtained using the information from Table A.1 in Appendix A. Lifetime earnings are calculated assuming a high school or lower educated Mexican resident works for 43 years and each year her earnings increase due to accumulated experience. The lifetime earnings is a present value measure obtained by discounting using yearly interest rate equal to 4%. The figure is also adjusted to reproduce a value in US dollars PPP adjusted.

dollars.³⁸ This figure is impressive if compared to actual Mexican earnings. Perhaps factors like proximity to colleges and the fact that some of the direct costs associated with attending a college are priced in US dollars contribute to explaining the relative high costs. However, the high cost of a college education in Mexico could also be explained by the inefficient financial sector that in Mexico implies much higher costs for borrowing to finance the period of study and this might make it very difficult for a large part of the Mexican population to actually access a college education. That is, Mexican students may face borrowing constraints that Americans do not face.

For the same reasons immigrants face higher educational costs than Americans. Their foregone earnings are closer to the foregone earnings of Americans since they have the option to migrate earlier and work in the US without a college education or wait until they are college educated and migrate later. However, the fact that they need to acquire their education in Mexico implies that they face the same difficulties that non-migrants face.

Table 5: Parameter Estimates

Parameter	Model 1			Model 2			Model 3	
	Point Est.	s.e.		Point Est.	s.e.		Point Est.	s.e.
$\pi_{mx,C}$	0.593	0.191		0.596	0.058		0.598	0.086
τ_{mx}	1.091	0.235		1.095	0.069		1.095	0.123
$\pi_{us,H}$	1.950	0.062		1.944	0.019		1.943	0.057
$\pi_{us,C}$	1.962	0.049		1.960	0.052		1.964	0.046
τ_m	0.554	0.113		0.258	0.333		0.260	0.122
τ_{us}	0.102	0.101		0.109	0.079		0.109	0.056
z	0.405	0.074	z_H	0.346	0.021	z_H	0.346	0.017
			z_C	0.642	0.249	z_C	0.644	0.060
ξ	0.006	0.227		0.003	0.063			
μ_ψ	0.000	0.000		2.927	0.020		0.000	0.000
σ_ψ^2	2.920	0.104		0.323	0.013		2.925	0.058
$\eta(\hat{\theta}, \hat{\gamma})$	147.88			9.43			11.86	

Rows 3 and 4 show the skill prices for Mexicans working in the US. The skill price for high school educated Mexicans working in the US ($\pi_{us,H}$) is 1.94, while for college educated Mexicans in the US ($\pi_{us,C}$) it is 1.96, both significantly different from zero at the 5% level.

³⁸The cost is calculated as $50,000 * (e^{\pi_{mx,C}} - e^{\pi_{mx,C} - \tau_{mx}})$. This cost is net of self selection in to college. Since the cost of attending college is proportional to the college earnings, the self selection implies an even higher cost for those who attend college.

The difference between the college and the high school skill price for Mexicans born in the US measures the return to college net of self selection, which in this case is about 2% and not significantly different from zero.³⁹ Row 5 in the table shows the cost of college attendance for Mexican immigrants (τ_m), which is estimated at 0.26, equivalent to about 81,000 US dollars, while row 6 shows the cost for US born Mexicans (τ_{us}), estimated at 0.11, equivalent to 37,000 US dollars.⁴⁰

Rows 7 to 10 present the estimates of the loss of human capital. Row 7 shows the direct loss faced by immigrants with a high school education (z_H) while row 8 shows this for a college educated immigrant (z_C); they are, respectively, 0.35 and 0.64.⁴¹ These estimates are both significantly different from zero at the 5% level, and they are also significantly different at the same confidence level.⁴² Therefore, the model suggests that the intellectual ability is more difficult to adapt to the US labor market than the manual ability. This explains why immigrants have lower returns from their education than other generations without resorting to explanations based on negative selection. In Model 3, the loss of capacity to transmit human capital to their children (ξ) is constrained to be zero, however, columns 1 and 2 show that the other models estimate this parameter to be close to zero. Finally, the estimates for the mean and the variance of the utility cost distribution are 2.92 and 0.32 respectively, both significant at the 5% level.⁴³

As mentioned in Section 5.1 the parameter β is not estimated, rather the value of 0.3079 is assigned to it. In order to evaluate the importance of this parameter I perform a sensitivity analysis and re-estimate the model with two alternative values: 0.154 - Model Beta-1 and

³⁹Although it is not reported here, Mayer (2008) also estimates this difference for Americans and reports a value of 4.8% with a standard error of 0.08. This suggests that the returns to college for Mexicans in the US and Americans are not significantly different.

⁴⁰The dollars figures are calculated as in footnote 38,

⁴¹Recalling footnote 14, and the fact that the direct loss of human capital is mainly identified by the earnings difference between Mexican immigrants and the second generation Mexicans for each educational group, it is possible to conclude that this parameter might be biased upward. In fact, if Mexican immigrants with high school education or lower had the same average years of schooling than their second generation counterpart, their earnings gap would have been lower, predicting a direct loss of human capital for this group of immigrants lower than the one estimated.

⁴²A formal test on their difference gives an estimate of $z_C - z_H = 0.29$ with a standard error of 0.071.

⁴³Other than the parameters z and ξ , the only other notable difference between Model 1 and 3 is the value that takes the cost of education for first generation Mexicans. In Model 1, it is estimated to be 0.55 compared to 0.26 in Model 3. The reason is that, allowing for different human capital losses, Model 3 estimates a much greater loss for college than high school educated immigrants, which is an implicit additional cost of education.

0 - Model Beta-2. Table 6 shows the re-estimated parameters together with the benchmark model. In general the parameters do not differ substantially from one model to another, with the notable exception of the variance of the disutility cost. The benchmark model estimates a variance of 0.32; the model with a β reduced by half estimates a variance of 0.14; and the model with no altruism, $\beta = 0$, estimates a variance of 0.04. Another important difference between the three models is the goodness of fit, which shows that both of the alternative models perform worse than the benchmark. However, the distance between the benchmark model and the first model is much lower than between the benchmark and the second model.

Table 6: Parameter Estimates

Parameter	Benchmark Model		Model Beta-1		Model Beta-2	
	Point Est.	s.e.	Point Est.	s.e.	Point Est.	s.e.
$\pi_{mx,C}$	0.598	0.086	0.584	0.090	0.583	0.082
τ_{mx}	1.095	0.123	1.084	0.064	1.093	0.068
$\pi_{us,H}$	1.943	0.057	1.942	0.023	1.943	0.016
$\pi_{us,C}$	1.964	0.046	1.967	0.034	1.968	0.027
τ_m	0.260	0.122	0.253	0.461	0.276	0.165
τ_{us}	0.109	0.056	0.110	0.016	0.110	0.026
z_H	0.346	0.017	0.345	0.032	0.317	0.015
z_C	0.644	0.060	0.640	0.369	0.579	0.175
μ_ψ	2.925	0.058	2.148	0.072	1.674	0.013
σ_ψ^2	0.323	0.023	0.140	0.050	0.038	0.012
$\eta(\hat{\theta}, \hat{\gamma})$	11.86		13.69		22.28	

The reason why the model with no altruism performs substantially worse than the other two is that it is not capable of reproducing the evolution of education and earnings from the second to the third generation, which identifies the selection mechanism.⁴⁴ The model with a lower but positive level of altruism is capable of fitting the evolution as much as the benchmark model, and does this by estimating a much lower variance for the disutility shock distribution. The altruism drives the positive selection, and the greater it is, the stronger is the positive selection. However, a stronger positive selection can also be obtained with a lower variance of the distribution of the disutility cost. Therefore, by changing β and σ_ψ^2 in the same direction,

⁴⁴Table B.2 in Appendix B shows the moments simulated under the three alternative models. The first two models do not differ substantially and this explains the close fit. The worse fit of the third model comes from two main moments the model is not capable to reproduce: the college share and the earnings of the second generation.

it is possible to keep constant the fit of the model for a large range of values for the two parameters. This explains why it is impossible to identify both parameters at the same time. However, if we set $\beta = 0$, the model cannot reproduce the positive selection necessary to fit the model for any value of the variance of the disutility cost, resulting in a worse fit of the model. Therefore, the empirical analysis of the model provides evidence of the importance of altruism in determining the migration decision, and in determining the positive self selection.

Overall the results from the estimation suggest four important facts: 1) immigrants face an important loss of human capital used in the labor market upon migration; 2) the loss of human capital for college educated immigrants is higher than for immigrants with high school education; 3) there is no loss of capacity to transfer human capital to children; and 4) altruism is an important factor in determining the migration decision and the selection mechanism. The first two facts imply that the human capital actually brought by immigrants is greater than what the earnings suggest. Previous studies focused on the earnings of immigrants to assess the amount of human capital they bring and the consequent selection pattern. Compared to previous studies, I find that the actual human capital brought by immigrants is 0.644 log-points higher conditional on college educated and 0.346 log-points higher for high school educated immigrants when the capacity to transfer it to future generations is considered. In fact, the third fact implies that, while the first generation incurs a significant and sizable loss of human capital, their children are not affected by this loss and can expect to inherit the same amount of abilities, whether or not the parents migrate. This result is also consistent with Caponi (2006) that demonstrates that the loss of human capital is a positive function of education. The second result explains why there is a jump in the earnings from the first to the second generation of Mexicans in the US, a jump that is not present from the second to the third generation. This explanation does not require negative self selection as postulated by Borjas (1993). The following section looks in more detail at how the model works with particular attention at how the self selection mechanism is determined.

6.1 Self Selection and Intergenerational Assimilation

A question that has important policy implications and that motivates a large part of the migration literature is how “good” are the immigrants entering the host country; where good

refers to how skilled they are, and how likely they are to successfully integrate. Using the estimated model we can evaluate the quality of Mexican immigrants in terms of unobservable as well as observable characteristics, and as such evaluate the amount of human capital brought to the US by the first generation and how this human capital is transferred to the successive generations. Table 7 reports the average abilities for Mexican immigrants and non-migrants in Mexico unconditional and conditional on the educational choice.

Table 7: Average Abilities and self selection

Decision	Manual Ability (H)	Intellectual Ability (C)
Migrant	0.025	0.049
Non-migrant	-0.003	-0.009
Migrant with High School	0.042	-0.051
Migrant with College	-0.088	0.711
Non-migrant with High School	0.016	-0.128
Non-migrant with College	-0.109	0.641

Rows 3 and 5 of Table 7 show the averages of the manual and intellectual ability conditional on high school for migrants and non-migrants, respectively. For the manual ability, the average ability for migrants is higher than for non-migrants, which indicates positive selection. The same is also true for the average intellectual ability, which indicates that high school educated immigrants are positively selected with respect to both abilities. A comparison of rows 4 and 6, reveals the selection conditional on choosing a college education. Again, immigrants are positively selected with respect to both abilities. Not surprisingly, the first two rows of the table show that, unconditional on the educational choice, the averages of both abilities are higher for immigrants than for non-migrants.⁴⁵ Therefore, immigrants are positively selected with respect to both the abilities.⁴⁶

⁴⁵Note that the unconditional value of the intellectual ability is negative in both cases. This is because the average intellectual ability in Mexico is affected by the positive selection of immigrants, or negative selection of non-migrants. Therefore, new generations of Mexicans have on average less intellectual ability than they would if there was not migration, *i.e.* zero.

⁴⁶It should be noted that the model assumes that fertility is independent on abilities and equal for every individual. Differential fertility across educational groups might have an impact on the average earnings on different generations. If lower ability parents tend to have more children than higher ability parents the decreased average earnings among college educated could be due in part to differential fertility. Yet Preston and Campbell (1993), for example, finds evidence that a worsening in the ability distribution from one generation to another cannot be explained uniquely by fertility differentials but must be accompanied by a higher than *steady state* equilibrium average ability in the original generation. Moreover, as Retherford and Sewell (1989) claim, the evidence of a relationship between ability and fertility is much weaker for men than for women. Therefore,

That immigrants are positively self selected with respect to the intellectual ability might seem counterintuitive since the college educated are the ones who are most affected by the human capital loss. Indeed it is the selection with respect to the intellectual ability that is strong and positive and that drives the positive selection of both abilities.⁴⁷ The selection with respect to the manual ability is negative but weaker. Due to the positive correlation between the two abilities, this implies that it is the intellectual ability that drives the selection mechanism. The reason for the positive selection is altruism. Immigrants with higher level of intellectual ability migrate for the better educational opportunities offered in the US to their children and grandchildren.

This finding is consistent with proposition 3 in Mayer (2008) and reinforces it. In his paper Mayer proves that a strong positive correlation between the two abilities creates a positive correlation between parents' earnings and the probability that children attend college. In the case of immigrants, parents with larger amounts of intellectual ability tend to migrate more and tend to choose to remain high school educated. However, they migrate with the expectation of their children becoming college educated. This suggests that even if less educated, immigrants are among the best individuals in terms of their levels of intellectual ability.

Table 8 shows that the intergenerational assimilation into the US labor market of successive generation of Mexicans should be faster than the earnings of the first generation suggest. The table reports the human capital, valued at market prices, by sector and aggregate in the benchmark case and in a counterfactual scenario where ξ is assumed to be equal to 1. That is, the loss of human capital faced by immigrants is a permanent intergenerational loss transferred to the future generations. The figures reported are in thousands of US dollars per person, and assume that the present value of lifetime earnings for high school educated non-migrant Mexicans is 50,000 US dollars.⁴⁸ The first row shows the human capital of non-migrant Mexicans, the second the first generation of immigrants and the third row the second generation.

looking at males only, even accounting for differential fertility cannot rule out a role for positive self selection of immigrants to explain the lower earnings of college educated third generation Mexicans compared to their second generation counterpart.

⁴⁷Appendix D gives a proof that the positive selection is due to altruism and is driven by the intellectual ability.

⁴⁸See footnote 37 for an explanation on how this figure is obtained.

Table 8: Human Capital Accounting

Generation	Benchmark $\xi = 0$			Counterfactual $\xi = 1$		
	(H)	(C)	(A)	(H)	(C)	(A)
Non-Mig.	50.00	172.85	68.99	50.16	167.70	68.78
1 st Gen.	253.65	383.88	270.81	253.35	387.21	268.90
2 nd Gen.	357.85	589.64	463.52	358.30	580.92	412.03

Compared to the benchmark case, in the counterfactual experiment the aggregate human capital per person of the second generation drops by 50,000 US dollars from 463,000 to 413,000 US dollars, more than 10% - row 3, columns 3 and 6. By sector the difference is less pronounced and only appears for the college educated that in the counterfactual scenario lose about 5,000 US dollars. Therefore, most of the loss of human capital is due to a drop in the share of college educated among the second generation. In contrast, the aggregate human capital per person of the first generation is not substantially different between the benchmark and the counterfactual. This suggests that, if the loss of human capital is a one time/one generation loss, the human capital transmitted to the second generation is higher. This confirms that, as reported in the previous section, immigrants bring more human capital than their earnings show. This excess of human capital of immigrants can be inferred by looking at the second generation immigrants.

7 Policy Evaluation

In this section I first look at the effects of a policy that aims at integrating immigrants faster in the host country, then I look at the effects of a policy aimed at increasing the educational attainment in Mexico or of Mexicans in the US by lowering the cost of a college education.⁴⁹ Examples of integration policies are programs that teach the official language to immigrants, or that help immigrants to adapt their skills to the local labor market.

I assume that integration policies can be translated into lower losses of human capital. I also take into account two different scenarios, one in which the programs target high school educated immigrants, and therefore reduces the loss of human capital only for that group,

⁴⁹The model presented in this paper is a partial equilibrium model. The migration flow and composition may also have an effect on the skill prices that the partial equilibrium specification is unable to capture. Therefore, some caution must be used in interpreting the results of the counterfactual experiments.

and another in which the college educated are targeted. I assume, conservatively, that the cost of the policy is equal to value of the human capital “recovered.” Alternatively, the policy can be thought as giving a subsidy to immigrants of a certain educational group to integrate their salary. To properly compare the alternative scenarios, I impose that each policy has the same aggregate cost.⁵⁰ Because the number of immigrants affected by the second policy is smaller, the amount spent for each individual is larger. Each policy implies that the loss of human capital of the targeted group is lower than the benchmark case. When the high school educated are targeted their loss of human capital decreases from 0.3457 to 0.2867; when the college educated immigrants are targeted their initial loss of 0.6444 becomes 0.4584.

Table 9 reports the human capital, valued at market prices, by sector and aggregate under the alternative policies and the benchmark. The figures reported are in thousands of US dollars per person, and assume that the present value of lifetime earnings for high school educated non-migrant Mexicans is 50,000 US dollars.⁵¹ Row 1, 2 and 3 show the human capital of non-migrant Mexicans, the first generation and the second generation, respectively. The last row shows the cost of the policy per person. The effectiveness of the policies can be compared by looking at the value of the aggregate human capital per person for the first and the second generation. Compared to the benchmark case, the aggregate human capital of the first generation increases from about 271,000 to 278,000 US dollars, an increase of less than 8,000 US dollars per person. This compares with a cost of the policy of about 12,600 US dollars per person. Overall, with the first policy we have a net loss. Moreover, the human capital per person for the second generation is lower under the first policy compared to the benchmark. Therefore, the policy does not bring any gain with future generations either.

The second policy compares much more favorably than the first to the benchmark. By targeting the college educated the human capital of the first generation increases from 270,000 to 315,000 US dollars per person, an increase of about 45,000 US dollars. Compared to the 13,280 US dollars of its cost, the improvement is much larger and implies a net gain.⁵²

⁵⁰I first assume that an equivalent of about 13,760 US dollars per person is spent to increase the human capital of immigrants with high school or less. Given the migration and education choices I calculate the total cost of the policy by multiplying the per-person amount by the total number of immigrants with high school or less, and impose that the second policy has the same aggregate cost.

⁵¹See footnote 37 for an explanation on how this figure is obtained.

⁵²The cost of the policy per person is higher under the second policy than the first because the migration rate is lower. Since the aggregate cost is kept constant, a lower number of immigrants increases the cost per person.

Table 9: Counterfactual Simulation: Policy Evaluation - Human Capital

Generation	Benchmark			Counterfactual 1			Counterfactual 2		
	(H)	(C)	(A)	(H)	(C)	(A)	(H)	(C)	(A)
Non-Mig.	50.00	172.85	68.99	50.16	167.70	71.91	49.98	172.55	68.57
1 st Gen.	253.65	383.88	270.81	268.73	387.21	280.98	254.14	456.59	285.48
2 nd Gen.	357.85	589.64	463.52	358.30	580.92	457.65	360.03	581.34	465.04
Pol. Cost				13.76		12.34		35.07	5.43
Counterfactual 1: $z_H = 0.2867$, $z_C = 0.6444$									
Counterfactual 2: $z_H = 0.3457$, $z_C = 0.4584$									

Moreover, the gain is transferred to the second generation as well, which earns 18,000 US dollars more under the second policy compared to the benchmark. Overall, by targeting college educated immigrants, the policy intensifies the brain drain from Mexico attracting more highly skilled immigrants. The effect of the brain drain can also be seen by looking at the non-migrant Mexicans in row 1. Under the first policy the aggregate human capital per person slightly increases compared to the benchmark, implying a mild reduction of the brain drain. Under the second policy there is a marked decline due to a stronger brain drain.

Overall the policy experiment suggests that policies aiming at rapidly integrating immigrants in the host country can have the effect of increasing the earnings potential of immigrants. However, a cost-benefit analysis shows that a policy that targets college educated immigrants is much more effective than one that targets high school educated. Targeting college educated not only benefits the first generation immigrants but also their children and grandchildren, who benefit from the stronger positive self selection.

The other two scenarios presented here consider an exogenous change to the cost of a college education in Mexico and in the US. The first scenario assumes that there is a 5% decrease in the cost of college in Mexico represented by a transfer of 3,000 dollars to every student attending college, the second assumes a decrease in the US cost of college by a transfer of the same amount to every student attending a college in the US.

Table 10 reports the human capital by sector and aggregate under the alternative policies and the benchmark. Columns 3 and 6 in second row show that decreasing the cost of attending college in Mexico reduces the average human capital brought by immigrants into the US.

Table B.5 in Appendix B reports the simulated moments under both policies.

The reduction is transferred to the second generation, which has a lower human capital per person. In contrast, Column 1 show that the human capital possessed by Mexicans increases significantly. A look at the first row of Table B.6 in Appendix B also shows that the overall migration rate declines slightly. Overall, by effect of the policy there would be a lower flow of migrants especially among the most educated and with higher intellectual ability. Looking at columns 2 and 4, in second row it is possible to notice that the average human capital for college educated actually increases among the immigrants compared to the benchmark case. This means that conditional on college education the selection mechanism is actually positive and stronger than in the benchmark. However, this fact is mainly due to the stronger selection into education among immigrants rather than the selection into migration for people with high intellectual ability. In fact, as row 3 clarifies, the unconditional average level of intellectual ability among immigrants is lower under the policy than in the benchmark case. This is why the average human capital among the second generation Mexicans with a college education is much lower under the policy than in the benchmark.⁵³ Overall, the effect of the policy is to contain the brain drain from Mexico to the US. Immigrants are less self-selected with respect to the intellectual ability, even though conditional on being college educated they are more selected than in the benchmark case. In contrast, the non-migrant Mexicans have a higher human capital per person. The migration rate falls when the cost of college college decreases in Mexico, and, while the share of college educated among non-migrant Mexicans increases, among first generation immigrants decreases. This fact explains the increase in the aggregate human capital per person in Mexico and its fall in the US.

Table 10: Counterfactual Simulation: Reducing the Cost of College - Human Capital									
Generation	Benchmark			Lower College Cost MX			Lower College Cost US		
	(H)	(C)	(A)	(H)	(C)	(A)	(H)	(C)	(A)
Non-Mig.	50.00	172.85	68.99	50.31	163.42	74.64	49.99	172.80	68.94
1 st Gen.	253.65	383.88	270.81	253.08	386.14	263.78	253.75	384.21	270.98
2 nd Gen.	357.85	589.64	463.52	359.31	570.91	451.86	359.66	581.84	463.27

The second scenario shows the effects of an decrease in the cost of a college education in

⁵³Moreover, from Table B.6 in Appendix B it is possible to see that even the educational attainment is slightly worse.

the US. In this case the results are opposed to the ones reported above. The human capital per person in Mexico decreases, while it increases in the US for the first and second generation. This is primarily due to the increase in the college share for the first and second generation, and a reduction for the non-migrants Mexicans.⁵³ Overall, the increased skill premium in the US drives a stronger positive selection with respect to the intellectual ability which mainly translates in more college educated among the first and the second generation Mexicans.

8 Conclusion

This paper provides an explanation of the intergenerational dynamics of earnings and educational attainment of successive generations of Mexicans in the US that does not rely on negative self selection. The explanation is based on three main concepts. The first concept is that immigrants have difficulties adapting their abilities in the host country. This includes language ability, social skills, and different cultural traits that represent the formidable challenge of adapting acquired skills from one's mother country to another country. This difficulty translates to a reduced capacity toward using one's abilities to produce earnings and, therefore, results in lower earnings. The second concept is that individuals are endowed with two abilities that can be used alternatively depending on the acquired level of education. The intellectual ability is used if some college education is acquired. Alternatively, the manual ability is used. The third concept is that there is a transfer of abilities from parents to their children. In this respect, immigrants' capacity to transfer their abilities to their children is not reduced. Therefore, while immigrants are observed to earn less because they find it difficult to adapt their skills to the host country, their children earn more because they can inherit all the abilities of their parents, including that part that could not be used for producing earnings.

A partial equilibrium intergenerational altruistic model that is capable of interpreting the main features of Mexican migration is built and estimated. By allowing agents to be endowed with two distinct abilities the model is capable of capturing the complexity of the selection mechanism. The estimation results highlight some important facts: 1) immigrants face an important loss of human capital upon migration; 2) the loss of human capital for college educated immigrants is higher than for immigrants with high school education or lower; 3) there is no loss of capacity to transfer human capital to children; 4) altruism is an important

factor that motivates migration; and 5) immigrants are overall positively self selected with respect to their abilities.

Interestingly the paper also finds that even among the high school educated, immigrants are positively selected with respect to the intellectual ability. This is because immigrants care about their children. Parents with larger amounts of intellectual ability tend to migrate more and tend to choose to remain high school educated. However, they migrate with the expectation of their children becoming college educated. Therefore, measures that rely on the earnings performance and educational attainment of immigrants underestimate the amount of human capital they bring into the host country.

In this sense, this paper reverses the pessimistic view implied by negative selection and intergenerational transmission of abilities theory proposed by Borjas (1993). A reason why new immigrant cohorts are observed to do worse in terms of earnings than the previous European based waves of immigrants may be given by a higher difficulty to adapt their skills to the new country. However, future generations of Mexican Americans should be observed to assimilate as fast as other previous ethnic groups did, provided that there are not other exogenous obstacles that prevent this integration.

Finally, the paper evaluates alternative policies that aim at integrating immigrants into the US labor market, or that contrast the brain drain from Mexico to the US. The simulation results show that the targeting college educated rather than high school educated immigrants generates larger effects on human capital. By reducing the loss of human capital faced by college educated immigrants, the policy generates direct positive effects on the overall human capital and earnings of the first generation, but also indirect effects on the first and the second generation by strengthening the self selection process of immigrants. In contrast, the policy that contrasts the brain drain by reducing the cost of education in Mexico. The simulation results indicate that this policy is effective in contrasting the brain drain from Mexico to the US. Immigrants become less self-selected with respect to human capital and among non-migrant Mexicans the average human capital increases. Moreover, the policy also reduces the migration rate especially among the highly educated Mexicans.

References

- ALTONJI, J. G., AND U. DORASZELSKI (2005): “The Role of Permanent Income and Demographics in Black-White Differences in Wealth,” *Journal of Human Resources*, XL(1), 1–30.
- ATTANASIO, O., AND C. BINELLI (2010): “Mexico in the 1990s: the Main Cross-Sectional Facts,” .
- BEHRMAN, J. R., AND M. R. ROSENZWEIG (2002): “Does Increasing Women’s Schooling Raise the Schooling of the Next Generation?,” *American Economic Review*, 92.
- BINELLI, C. (2009): “The demand-supply-demand twist: how the wage structure got more convex,” *Mimeo, Oxford University*.
- BORJAS, G. J. (1987): “Self Selection and the Earnings of Immigrants,” *American Economic Review*, 77(4), 531–553.
- (1993): “The Intergenerational Mobility of Immigrants,” *Journal of Labor Economics*, 11(1), 113–135.
- (2000): “The Economic Progress of Immigrants,” in *Issues in the Economics of Immigration*, ed. by G. J. Borjas. University of Chicago Press, Chicago.
- BRIGHAM, C. C. (1923): *A Study of American Intelligence*. Princeton University Press, Princeton, NJ.
- CAPONI, V. (2006): “Heterogeneous Human Capital and Migration: Who Migrates from Mexico to the US?,” *IZA - Institute for the Study of Labor, Bonn, Germany*, (Discussion Paper N.2446).
- CARD, D. (2005): “Is the New Immigration Really so Bad?,” *Economic Journal*, 115(507), F300–F323.
- CHIQUEAR, D., AND G. H. HANSON (2005): “International Migration, Self Selection, and the Distribution of Wages: Evidence from Mexico and United States,” *Journal of Political Economy*, 113(2), 239–281.

- CHISWICK, B. R. (1978): "The Effects of Americanization on the Earnings of Foreign-Born Men," *Journal of Political Economy*.
- CUNHA, F., AND J. HECKMAN (2007): "The Technology of Skill Formation," *American Economic Review*, 97(2), 31–47.
- DUNCAN, B., AND S. J. TREJO (2005): "Ethnic Identification, Intermarriage, and Unmeasured Progress by Mexican Americans," *IZA - Institute for the Study of Labor, Bonn, Germany*, (Discussion Paper N.1629).
- GLICK, J. E., AND M. J. WHITE (2004): "Post-secondary school participation of immigrant and native youth: the role of familial resources and educational expectations," *Social Science Research*, 33(2), 272–299.
- GOURINCHAS, P.-O., AND J. A. PARKER (2002): "Consumption over the Life Cycle," *Econometrica*, 70(1), 47–89.
- HECKMAN, J., AND J. SCHEINKMAN (1987): "The Importance of Bundling in a Gorman-Lancaster Model of Earnings," *The Review of Economic Studies*, 54(2), 243–255.
- HECKMAN, J. J., AND G. L. SEDLACEK (1990): "Self-Selection and the Distribution of Hourly Wages," *Journal of Labor Economics*, 8(1), S329–S363.
- HECKMAN, J. J., J. STIXRUD, AND S. URZUA (2006): "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," NBER Working Papers 12006, National Bureau of Economic Research, Inc.
- KAO, G., AND M. TIENDA (1995): "Optimism and Achievement: The Educational Performance of Immigrant Youth," *Social Science Quarterly*, 76.
- KUROKAWA, Y. (2006): "Skill Intensity Reversal and the Rising Skill Premium: Evidence from the U.S. and Mexico," MPRA Paper 14013, University Library of Munich, Germany.
- LAIBSON, D., A. REPETTO, AND J. TOBACMAN (2005): "Estimating Discount Functions with Consumption Choices over the Lifecycle," *Mimeo Harvard*.

- MAYER, A. (2008): “Education, Self-Selection, and Intergenerational Transmission of Abilities,” *Journal of Human Capital*, 2(1), 106–128.
- McFADDEN, D. (1989): “A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration,” *Econometrica*, 57(5), 995–1026.
- McKEEVER, M., AND S. L. KLINEBERG (1999): “Generational Differences in Attitudes and Socioeconomic Status among Hispanics in Houston,” *Sociological Inquiry*, 69.
- NEWKEY, W. K., AND D. McFADDEN (1994): *Large Sample Estimation and Hypothesis Testing*, 2111–2245, Handbook of econometrics. Volume 4. Elsevier, North-Holland, Handbooks in Economics, vol. 2. Amsterdam; London and New York.
- PRESTON, S. H., AND C. CAMPBELL (1993): “Differential Fertility and the Distribution of Traits: The Case of IQ,” *The American Journal of Sociology*, 98(5).
- RETFERFORD, R. D., AND W. SEWELL (1989): “How Intelligence Affects Fertility,” *Intelligence*, 13, 169–185.
- SATTINGER, M. (1993): “Assignment Models of the Distribution of Earnings,” *Journal of Economic Literature*, 31(2), 831–80.
- WILLIS, R. J., AND S. ROSEN (1979): “Education and Self-Selection,” *Journal of Political Economy*, 87(5), S7–36.

Appendix A Data, Moments and Covariance Matrix

This section illustrates how the moments are created from the data and how the covariance matrix is derived. The data set is formed by pooling together observations on Mexicans living in Mexico, and Mexicans living in the US, of first, second and third generation. Information about Mexican born individuals living in the US is obtained using the pooled 1994-2008 CPS data. In order to obtain an estimate of the Mexican population living in the US in 2000, I re-weight the CPS observations in each survey year different from 2000 to obtain an aggregate number of the Hispanic population equal to the number present in 2000. Then I divide the weight of all observations by the number of surveys used. The re-weighting guarantees that the sum of all weighted observations from the pooled CPS data reproduce the Mexican population present in the US in 2000. Information on Mexicans living in Mexico is obtained using the 2000 Mexican census. I use a public use micro sample of 1% of the Mexican population in order to obtain an estimate of the total Mexican male population between 22 and 75 years old. I then pool the Mexican census with the CPS data to obtain one single data set containing all the information on Mexicans living in Mexico and in the US. Once I have the unified data set with the corrected weights, I select all male individuals between 22 and 75 who have positive earnings, work full time, do not attend school and, in case of first generation immigrants, who entered the US at an age of 22 or older. Earnings are corrected to take into account top-coded values, and the sample excludes outliers in terms of hourly wages.^{A.1,A.2}

The moments are based on the following regression model,

^{A.1}To build a measure of log hourly wage I use observations on yearly income from the CPS. These observations are top-coded at different levels depending on the survey year. In 1994 and 1995 incomes over 100,000 dollars were top-coded. From 1996 to 2002 the level was 150,000, and then increased to 200,000 since 2003. From 1996 the CPS does not set all the top-coded observations equal to the top-code level. Instead the averages incomes of six categories of individuals conditional on being top-coded are calculated. These categories are Hispanics, blacks and whites divided by men and women. Then each top-coded observation is replaced with the conditional mean corresponding to the group of the individual with top-coded income. To correct for top-coding I first re-assign the top-coding threshold value to each top-coded observation, then I build a measure of log hourly earnings for all top-coded and non top-coded observations. Once I have this measure I calculate the expected mean value of the top-coded observations by estimating a Tobit model, assuming that log-hourly wages are normally distributed. Once I have the mean value I adjust each top-coded observation by the difference between the expected mean from the Tobit estimation and the top-coded value.

^{A.2}For the US earnings below half of the federal minimum wage in 2000 - \$5.15/2, and above \$250 are excluded. For Mexico wages below \$0.28 and above \$27.5, corresponding roughly to the extreme (left and right) 0.05% of the wage distribution.

$$\log(w_i) = \beta_0 + \beta_1 G_{1Hi} + \beta_2 G_{2Hi} + \beta_3 G_{3Hi} + \beta_4 G_{0Ci} + \beta_5 G_{1Ci} + \beta_6 G_{2Ci} + \beta_7 G_{3Ci} + \gamma X_i + \zeta_i \quad (\text{A.1})$$

where $\log(w_i)$ represents log-hourly wages regressed on a set of dummy variables G' s indicating different generation/education groups and a set individual characteristics X_i . Among the individual characteristics I use a quadratic function of age - centered at 48 years - interacted with the set of the G dummies to allow different effects for each group, geographical dummies and, for data on Mexicans in the US, the year of birth of the individual.^{A.3} Dummy variables were constructed to represent the generations of Mexicans: G_{iH} and G_{iC} with $i = 1, 2, 3$ for first, second and third generation with high school and college respectively, and G_{0C} for Mexicans living in Mexico with college education. The reference group captured by the intercept refers to Mexicans living in Mexico with high school education.

Table A.1 reports the earnings moments resulting from the estimation of equation (A.1). Focusing on the first column of Table A.1, the intercept coefficient in row 1 gives the average log hourly wage in 2000 US dollars of the high school or lower educated group of Mexican men residents in Mexico, which is also the group that has the lowest earnings. In row 5 an impressive 1.22 log-points, shows the returns to college in Mexico. The earnings of Mexican immigrants with a high school degree or less are on average 1.51 log-points higher than Mexicans in Mexico with the same level of education (row 2), while the second and third generations improve their earnings even further earning 1.97 and 1.96 log-points more (rows 3 and 4). A similar pattern is observed for college educated Mexicans living in the US. The first generation of immigrants earn on average 2.01 log-points more than high school educated Mexicans in Mexico. By the second generation the difference with the reference groups increases to 2.43 log-points, but then the earnings decrease for the third generation to about 2.36.

An earnings variable, net of the effect of other exogenous variables, is created as follows,

$$v_i = \log(w_i) - \hat{\gamma} X_i \quad (\text{A.2})$$

and form another set of dummy variables for generation groups as follows, D_{0H} for Mexicans living in Mexico with high school education, $D_0 = D_{0H} + D_{0C}$ for all Mexicans living in Mexico,

^{A.3}The geographical dummies are for states in Mexico and metropolitan status in the US. The reference groups are large cities for the US and the district of Mexico City for Mexico.

Table A.1: Earnings Gaps Between Different Generations of Mexicans

Dependent Var.: Log Hourly Wage	Standard Definition		Stricter Definition	
	Men	Women	Men	Women
Intercept	0.6474 (0.0129)	0.6039 (0.0197)	0.6444 (0.0124)	0.6105 (0.0191)
1 st Gen. Imm. w/o Coll.	1.6263 (0.0194)	1.3507 (0.0274)	1.6231 (0.0167)	1.3629 (0.0249)
2 nd Gen. Mexicans w/o Coll.	1.9742 (0.0261)	1.6701 (0.0332)	1.9676 (0.0246)	1.6769 (0.0322)
3 rd Gen. Mexicans w/o Coll.	1.9642 (0.0209)	1.6384 (0.0272)	1.9636 (0.0188)	1.6504 (0.0254)
Mexicans with Coll.	1.2230 (0.0081)	1.1440 (0.0118)	1.2232 (0.0079)	1.1440 (0.0116)
1 st Gen. Imm. Coll.	2.0095 (0.0276)	1.7161 (0.0391)	2.0490 (0.0236)	1.7552 (0.0342)
2 nd Gen. Mexicans Coll.	2.4288 (0.0274)	2.1212 (0.0332)	2.4065 (0.0268)	2.1275 (0.0331)
3 rd Gen. Mexicans Coll.	2.3636 (0.0218)	2.0889 (0.0276)	2.3628 (0.0197)	2.1015 (0.0258)
N.OBS.	133567	64784	147044	69988
R^2	0.6424	0.6192	0.6689	0.6315

Standard errors in parenthesis.

$D_i = D_{iH} + D_{iC}$ with $i = 1, 2, 3$ for all first, second and third generation immigrants. Finally $D_{[01]} = D_0 + D_1$ for all Mexican born individuals. All the moments I use are calculated using regression models. The first set of moments are the migration rate and the share of college graduates among generations 0, 1, 2 and 3. The migration rate is obtained as the fraction of observations belonging to generation 1 over the total Mexican born individuals. I assume that the following model generates the data,

$$D_{1i} = R_m D_{[01]} + \epsilon_{mi} \quad (\text{A.3})$$

Similarly, to obtain the share of college educated individuals for each generation I assume the following models generate the data,

$$D_{jHi} = R_{ji} D_{ji} + \epsilon_{ji} \quad \text{for } j = 0, 1, 2, 3 \quad i = 1, \dots, N \quad (\text{A.4})$$

The second set of moments are the first moments of the earnings distributions conditional on the education/generation group. I obtain these moments assuming the following models,

$$v_i = \bar{v}_{jki}D_{jki} + \eta_{jki} \quad \text{for } j = 0, 1, 2, 3 \quad k = L, H \quad i = 1, \dots, N \quad (\text{A.5})$$

Assumptions about the error terms

The basic assumption about the error terms is that they are IID within and across each model. In other words, errors can be correlated across models only if they come from the same draw, or observation.

Variances and Covariances

The variance of an estimator of the first set can be easily derived by noticing first that,

$$\hat{R}_m = \frac{\sum D_1 \cdot D_{[01]}}{\sum D_{[01]}^2} = \frac{\sum (R_m D_{[01]} + \epsilon_m) \cdot D_{[01]}}{\sum D_{[01]}^2} = \frac{\sum R_m D_{[01]}^2 + \sum D_{[01]} \epsilon_m}{\sum D_{[01]}^2} = R_m + \frac{\sum D_{[01]} \epsilon_m}{\sum D_{[01]}^2} \quad (\text{A.6})$$

or, since $D^2 = D$,

$$\hat{R}_m - R_m = \frac{\sum D_{[01]} \epsilon_m}{\sum D_{[01]}} = \frac{\sum D_{[01]} \epsilon_m}{N_{[01]}} \quad (\text{A.7})$$

it follows that,

$$E[(\hat{R}_m - R_m)^2] = E\left[\frac{\left(\sum D_{[01]} \epsilon_m\right)^2}{N_{[01]}^2}\right] = \frac{E\left[\left(\sum D_{[01]} \epsilon_m\right)^2\right]}{N_{[01]}^2} \quad (\text{A.8})$$

and assuming there is no heteroscedasticity and that cross-correlations are all zeros, we have,

$$E[(\hat{R}_m - R_m)^2] = \frac{\sum D_{[01]} E(\epsilon_m^2)}{N_{[01]}^2} = \frac{\sum D_{[01]} \sigma_{\epsilon_m}^2}{N_{[01]}^2} = \frac{N_{[01]} \sigma_{\epsilon_m}^2}{N_{[01]}^2} = \frac{\sigma_{\epsilon_m}^2}{N_{[01]}} \quad (\text{A.9})$$

Similarly, we have that the share of college educated over the total population of generation zero is given by,

$$\hat{R}_{0H} = \frac{\sum D_{0H} \cdot D_0}{\sum D_0^2} = R_{0H} + \frac{\sum D_0 \epsilon_{0H}}{\sum D_0^2} \quad (\text{A.10})$$

or,

$$\hat{R}_{0H} - R_{0H} = \frac{\sum D_0 \epsilon_{0H}}{N_0} \quad (\text{A.11})$$

Therefore if we wanted to calculate the covariance between the migration rate and the college share of generation 0 we have,

$$E[(\hat{R}_m - R_m)(\hat{R}_{0H} - R_{0H})] = E\left[\frac{\sum D_{[01]}\epsilon_m}{N_{[01]}} \frac{\sum D_0\epsilon_{0H}}{N_0}\right] = E\left[\frac{\sum D_{[01]}D_0\epsilon_m\epsilon_{0H}}{N_{[01]}N_0}\right] \quad (\text{A.12})$$

or, since $D_{[01]}D_0 = D_0$

$$E[(\hat{R}_m - R_m)(\hat{R}_{0H} - R_{0H})] = \frac{\sum D_0 E[\epsilon_m\epsilon_{0H}]}{N_{[01]}N_0} = \frac{N_0 E[\epsilon_m\epsilon_{0H}]}{N_{[01]}N_0} = \frac{E[\epsilon_m\epsilon_{0H}]}{N_{[01]}} \quad (\text{A.13})$$

similarly,

$$E[(\hat{R}_m - R_m)(\hat{R}_{1H} - R_{1H})] = \frac{\sum D_1 E[\epsilon_m\epsilon_{1H}]}{N_{[01]}N_1} = \frac{N_1 E[\epsilon_m\epsilon_{1H}]}{N_{[01]}N_1} = \frac{E[\epsilon_m\epsilon_{1H}]}{N_{[01]}} \quad (\text{A.14})$$

Other covariances for the first set of moments are all zero, as it easy to verify,

$$E[(\hat{R}_{iH} - R_{iH})(\hat{R}_{jH} - R_{jH})] = E\left[\frac{\sum D_i D_j \epsilon_{iH} \epsilon_{jH}}{N_i N_j}\right] = \frac{\sum D_i D_j E[\epsilon_{iH} \epsilon_{jH}]}{N_i N_j} \quad (\text{A.15})$$

and since when $i \neq j$ $D_i D_j = 0$, the covariance is zero.

The second set of moments is given by the earnings. Similarly to before we can write,

$$\hat{v}_{jk} = \frac{\sum v \cdot D_{jk}}{\sum D_{jk}^2} = \bar{v}_{jk} + \frac{\sum D_{jk} \eta_{jk}}{N_{jk}} \quad (\text{A.16})$$

and therefore,

$$\hat{v}_{jk} - \bar{v}_{jk} = \frac{\sum D_{jk} \eta_{jk}}{N_{jk}} \quad (\text{A.17})$$

Because whenever $jk \neq xy$, $D_{jk}D_{xy} = 0$ all the covariances between the moments belonging to the second set of moments are zero. More interesting is to look at the covariances between moments of the first and the second set. Let's start with the migration rate,

$$E[(\hat{R}_m - R_m)(\hat{v}_{jk} - \bar{v}_{jk})] = E\left[\frac{\sum D_{[01]}\epsilon_m}{N_{[01]}} \frac{\sum D_{jk}\eta_{jk}}{N_{jk}}\right] = \frac{\sum D_{[01]}D_{jk}E[\epsilon_m\eta_{jk}]}{N_{[01]}N_{jk}} \quad (\text{A.18})$$

for $j = 0, 1$, $D_{[01]}D_{jk} = D_{jk}$, therefore,

$$E[(\hat{R}_m - R_m)(\hat{v}_{jk} - \bar{v}_{jk})] = \frac{E[\epsilon_m\eta_{jk}]}{N_{[01]}} \quad (\text{A.19})$$

for $j = 2, 3$ the covariance is instead zero. For the other covariances,

$$E[(\hat{R}_{iH} - R_{iH})(\hat{v}_{jk} - \bar{v}_{jk})] = \frac{\sum D_i D_{jk} E[\epsilon_i \eta_{jk}]}{N_i N_{jk}} \quad (\text{A.20})$$

if $i = j$ then the covariance is,

$$E[(\hat{R}_{jH} - R_{jH})(\hat{v}_{jk} - \bar{v}_{jk})] = \frac{E[\epsilon_j \eta_{jk}]}{N_j} \quad (\text{A.21})$$

otherwise is zero.

The following table recaps the covariance matrix among the above moments,

Table A.2: Original Covariance Matrix

Moments	\hat{R}_m	\hat{R}_{0H}	\hat{R}_{1H}	\hat{R}_{2H}	\hat{R}_{3H}	\hat{v}_{0L}	\hat{v}_{0H}	\hat{v}_{1L}	\hat{v}_{1H}	\hat{v}_{2L}	\hat{v}_{2H}	\hat{v}_{3L}	\hat{v}_{3H}
R_m	$\frac{\sum_N w_i \epsilon_{mi}^2}{N_{01} \sum_N w_i}$	$\frac{\sum_N w_i \epsilon_{mi} \epsilon_{0Hi}}{N_{01} \sum_N w_i}$	$\frac{\sum_N w_i \epsilon_{mi} \epsilon_{1Hi}}{N_{01} \sum_N w_i}$	0	0	$\frac{\sum_N w_i \epsilon_{mi} \hat{\eta}_{0Li}}{N_{01} \sum_N w_i}$	$\frac{\sum_N w_i \epsilon_{mi} \hat{\eta}_{0Hi}}{N_{01} \sum_N w_i}$	$\frac{\sum_N w_i \epsilon_{mi} \hat{\eta}_{1Li}}{N_{01} \sum_N w_i}$	$\frac{\sum_N w_i \epsilon_{mi} \hat{\eta}_{1Hi}}{N_{01} \sum_N w_i}$	0	0	0	0
R_{0H}	$\frac{\sum_N w_i \epsilon_{0Hi}^2}{N_0 \sum_N w_i}$	$\frac{\sum_N w_i \epsilon_{0Hi} \hat{\eta}_{0Li}}{N_0 \sum_N w_i}$	0	0	0	$\frac{\sum_N w_i \epsilon_{0Hi} \hat{\eta}_{0Hi}}{N_0 \sum_N w_i}$	$\frac{\sum_N w_i \epsilon_{0Hi} \hat{\eta}_{1Hi}}{N_0 \sum_N w_i}$	0	0	0	0	0	0
R_{1H}	$\frac{\sum_N w_i \epsilon_{1Hi}^2}{N_1 \sum_N w_i}$	$\frac{\sum_N w_i \epsilon_{1Hi} \hat{\eta}_{0Li}}{N_1 \sum_N w_i}$	$\frac{\sum_N w_i \epsilon_{1Hi} \hat{\eta}_{1Li}}{N_1 \sum_N w_i}$	0	0	$\frac{\sum_N w_i \epsilon_{1Hi} \hat{\eta}_{0Hi}}{N_1 \sum_N w_i}$	$\frac{\sum_N w_i \epsilon_{1Hi} \hat{\eta}_{1Hi}}{N_1 \sum_N w_i}$	$\frac{\sum_N w_i \epsilon_{2Hi} \hat{\eta}_{2Li}}{N_2 \sum_N w_i}$	$\frac{\sum_N w_i \epsilon_{2Hi} \hat{\eta}_{2Hi}}{N_2 \sum_N w_i}$	0	0	0	0
R_{2H}	$\frac{\sum_N w_i \epsilon_{2Hi}^2}{N_2 \sum_N w_i}$	$\frac{\sum_N w_i \epsilon_{2Hi} \hat{\eta}_{0Li}}{N_2 \sum_N w_i}$	$\frac{\sum_N w_i \epsilon_{2Hi} \hat{\eta}_{1Li}}{N_2 \sum_N w_i}$	$\frac{\sum_N w_i \epsilon_{2Hi}^2}{N_2 \sum_N w_i}$	0	0	0	0	0	$\frac{\sum_N w_i \epsilon_{3Hi} \hat{\eta}_{3Li}}{N_3 \sum_N w_i}$	$\frac{\sum_N w_i \epsilon_{3Hi} \hat{\eta}_{3Hi}}{N_3 \sum_N w_i}$	0	0
R_{3H}	$\frac{\sum_N w_i \epsilon_{3Hi}^2}{N_3 \sum_N w_i}$	$\frac{\sum_N w_i \epsilon_{3Hi} \hat{\eta}_{0Li}}{N_3 \sum_N w_i}$	$\frac{\sum_N w_i \epsilon_{3Hi} \hat{\eta}_{1Li}}{N_3 \sum_N w_i}$	$\frac{\sum_N w_i \epsilon_{3Hi} \hat{\eta}_{2Li}}{N_3 \sum_N w_i}$	$\frac{\sum_N w_i \epsilon_{3Hi}^2}{N_3 \sum_N w_i}$	0	0	0	0	0	0	$\frac{\sum_N w_i \epsilon_{3Hi} \hat{\eta}_{3Li}}{N_3 \sum_N w_i}$	$\frac{\sum_N w_i \epsilon_{3Hi} \hat{\eta}_{3Hi}}{N_3 \sum_N w_i}$
\hat{v}_{0L}						$\frac{\sum_N w_i \hat{\eta}_{0Li}^2}{N_{0L} \sum_N w_i}$	0	0	0	0	0	0	0
\hat{v}_{0H}						$\frac{\sum_N w_i \hat{\eta}_{0Hi}^2}{N_{0H} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{0Li} \hat{\eta}_{0Hi}}{N_{0H} \sum_N w_i}$	0	0	0	0	0	0
\hat{v}_{1L}						$\frac{\sum_N w_i \hat{\eta}_{1Li}^2}{N_{1L} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{1Hi}^2}{N_{1H} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{1Li} \hat{\eta}_{1Hi}}{N_{1L} \sum_N w_i}$	0	0	0	0	0
\hat{v}_{1H}						$\frac{\sum_N w_i \hat{\eta}_{1Li}^2}{N_{1L} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{1Hi}^2}{N_{1H} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{1Li} \hat{\eta}_{1Hi}}{N_{1L} \sum_N w_i}$	0	0	0	0	0
\hat{v}_{2L}						$\frac{\sum_N w_i \hat{\eta}_{2Li}^2}{N_{2L} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{2Hi}^2}{N_{2H} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{2Li} \hat{\eta}_{2Hi}}{N_{2L} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{2Li} \hat{\eta}_{3Li}}{N_{2L} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{2Li} \hat{\eta}_{3Hi}}{N_{2L} \sum_N w_i}$	0	0	0
\hat{v}_{2H}						$\frac{\sum_N w_i \hat{\eta}_{2Li}^2}{N_{2L} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{2Hi}^2}{N_{2H} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{2Li} \hat{\eta}_{2Hi}}{N_{2L} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{2Li} \hat{\eta}_{3Li}}{N_{2L} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{2Li} \hat{\eta}_{3Hi}}{N_{2L} \sum_N w_i}$	0	0	0
\hat{v}_{3L}						$\frac{\sum_N w_i \hat{\eta}_{3Li}^2}{N_{3L} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{3Hi}^2}{N_{3H} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{3Li} \hat{\eta}_{3Hi}}{N_{3L} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{3Li} \hat{\eta}_{3Hi}}{N_{3L} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{3Li} \hat{\eta}_{3Hi}}{N_{3L} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{3Li} \hat{\eta}_{3Hi}}{N_{3L} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{3Li} \hat{\eta}_{3Hi}}{N_{3L} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{3Li} \hat{\eta}_{3Hi}}{N_{3L} \sum_N w_i}$
\hat{v}_{3H}						$\frac{\sum_N w_i \hat{\eta}_{3Li}^2}{N_{3L} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{3Hi}^2}{N_{3H} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{3Li} \hat{\eta}_{3Hi}}{N_{3L} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{3Li} \hat{\eta}_{3Hi}}{N_{3L} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{3Li} \hat{\eta}_{3Hi}}{N_{3L} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{3Li} \hat{\eta}_{3Hi}}{N_{3L} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{3Li} \hat{\eta}_{3Hi}}{N_{3L} \sum_N w_i}$	$\frac{\sum_N w_i \hat{\eta}_{3Li} \hat{\eta}_{3Hi}}{N_{3L} \sum_N w_i}$

Normalization

One additional problem is the normalization that is done to the moments by subtracting the lowest category average earnings to all the other earnings such that the earnings moments become,

$$\hat{v}_{jk} - \hat{v}_{0L} \quad \forall jk \neq 0L \quad (\text{A.22})$$

We then have seven earning moments instead of eight. In this case, the variances of the earning moments must take into account the normalization, so that, since the covariance between each original earning average is zero we can write,

$$V(\hat{v}_{jk} - \hat{v}_{0L}) = V(\hat{v}_{jk}) + V(\hat{v}_{0L}) \quad (\text{A.23})$$

Moreover, the covariances between the new earning moments need also to be re-evaluated,

$$Cov(\hat{v}_{jk} - \hat{v}_{0L}, \hat{v}_{xy} - \hat{v}_{0L}) = E[(\hat{v}_{jk} - \hat{v}_{0L} - \bar{v}_{jk} + \bar{v}_{0L})(\hat{v}_{xy} - \hat{v}_{0L} - \bar{v}_{xy} + \bar{v}_{0L})] \quad (\text{A.24})$$

Once again, because all the covariances between the original earning moments are zero, this simplifies to,

$$Cov(\hat{v}_{jk} - \hat{v}_{0L}, \hat{v}_{xy} - \hat{v}_{0L}) = V(\hat{v}_{0L}) \quad (\text{A.25})$$

In all the other covariances involving earnings moments, the covariance between the average earning of the lowest earning group and the moment considered must be subtracted. For example,

$$Cov(\hat{R}_{jH}, \hat{v}_{jk} - \hat{v}_{0L}) = Cov(\hat{R}_{jH}, \hat{v}_{jk}) - Cov(\hat{R}_{jH}, \hat{v}_{0L}) \quad (\text{A.26})$$

Appendix B Model Simulations

By looking at rows 6 and 10 in table B.1, it is possible to understand why Models 2 and 3 perform substantially better than Model 1. In Model 1, the simulated earnings of high school educated immigrants (row 6) are much lower than the data counterpart, while the earnings of college educated (row 10) are higher. In Models 2 and 3, these differences vanish almost completely.

Table B.1: Data Moments for the SMM Estimation: Males only

Moment	Data	Sim. Model 1	Sim. Model 2	Sim. Model 3
Migration Rate	0.100	0.101	0.099	0.100
College Mex. in Mex.	0.154	0.155	0.154	0.155
College 1 st Gen. in US	0.134	0.129	0.133	0.132
College 2 st Gen. in US	0.454	0.454	0.453	0.456
College 3 st Gen. in US	0.437	0.436	0.434	0.437
Earnings HS 1 st Gen. in US	1.626	1.569	1.624	1.624
Earnings HS 2 st Gen. in US	1.974	1.975	1.972	1.968
Earnings HS 3 st Gen. in US	1.964	1.970	1.964	1.963
Earnings C. Mex. in Mex.	1.223	1.219	1.222	1.222
Earnings C. 1 st Gen. in US	2.009	2.251	2.013	2.015
Earnings C. 2 st Gen. in US	2.429	2.415	2.411	2.423
Earnings C. 3 st Gen. in US	2.364	2.403	2.403	2.404

B.1 Sensitivity Analysis: β

Table B.2 shows the moments simulated under the three alternative models. The first two models do not differ substantially and this explains the close fit. The worse fit of the third model comes from two main moments the model is not capable to reproduce: the college share and the earnings of the second generation.

Table B.3 reports the self selection patterns under the alternative models. The table clearly shows that the model with no altruism predicts negative instead than positive selection with respect to the intellectual ability. A small positive selection with respect to the manual ability persists even with no altruism. The self selection with no altruism is entirely driven by those potential migrants that switch education if they migrate. If their disutility cost is too high to migrate they stay in Mexico and acquire a college education, if it is low enough, then migrate and remain with a high school education. The probability to migrate for “switchers”

Table B.2: Counterfactual Experiment

Moment	Benchmark	Counterf. 1 ($\beta = 0.1540$)	Counterf. 2 ($\beta = 0$)
Migration Rate	0.100	0.101	0.098
College Mex. in Mex.	0.155	0.154	0.154
College 1 st Gen. in US	0.132	0.132	0.131
College 2 st Gen. in US	0.456	0.454	0.443
College 3 st Gen. in US	0.437	0.440	0.440
Earnings HS 1 st Gen. in US	1.624	1.625	1.628
Earnings HS 2 st Gen. in US	1.968	1.970	1.965
Earnings HS 3 st Gen. in US	1.963	1.962	1.961
Earnings C. Mex. in Mex.	1.222	1.211	1.221
Earnings C. 1 st Gen. in US	2.015	2.020	2.025
Earnings C. 2 st Gen. in US	2.423	2.421	2.396
Earnings C. 3 st Gen. in US	2.404	2.405	2.403

is increasing with the manual ability and decreasing with the intellectual one, which justifies the selection pattern.

Table B.3: Average Abilities and self selection

Decision	Counterfactual 1 ($\beta = .158$)		Counterfactual 2 ($\beta = 0$)	
	Ability (H)	Ability (C)	Ability (H)	Ability (C)
Migrant	0.028	0.049	0.004	-0.007
Non-migrant	-0.004	-0.009	-0.000	0.006
Migrant with High School	0.044	-0.052	0.021	-0.107
Migrant with College	-0.079	0.709	-0.109	0.655
Non-migrant with High School	0.016	-0.128	0.019	-0.113
Non-migrant with College	-0.110	0.642	-0.106	0.657

Table B.4 reports the schooling choices of children conditional on the education of parents. It possible to see that the worsening in the educational distribution is driven primarily by more intergenerational persistence in the lower educational category and consequently a lower advancement in the educational attainment from one generation to the next.

Table B.4: Children Education Conditional on Parents Education

Parent/Child	Benchmark	Counterfactual 1 ($\beta = .158$)	Counterfactual 2 ($\beta = 0$)
<i>HS/HS</i>	0.507	0.507	0.518
<i>HS/Coll.</i>	0.362	0.360	0.351
<i>Coll./HS</i>	0.037	0.039	0.039
<i>Coll./Coll.</i>	0.094	0.094	0.093

B.2 Policy Evaluations: Simulated Moments

Table B.5: Counterfactual Simulation: Policy Evaluation - Moments

Moment	Benchmark	Counterfactual 1	Counterfactual 2
		($z_H = 0.2867$, $z_C = 0.6444$)	($z_H = 0.3457$, $z_C = 0.4584$)
Migration Rate	0.100	0.094	0.105
College Mex. in Mex.	0.155	0.185	0.152
College 1 st Gen. in US	0.132	0.103	0.155
College 2 st Gen. in US	0.456	0.446	0.474
College 3 st Gen. in US	0.437	0.437	0.451
Earnings HS 1 st Gen. in US	1.624	1.619	1.626
Earnings HS 2 st Gen. in US	1.968	1.966	1.975
Earnings HS 3 st Gen. in US	1.963	1.960	1.965
Earnings C. Mex. in Mex.	1.222	1.188	1.221
Earnings C. 1 st Gen. in US	2.015	2.019	2.002
Earnings C. 2 st Gen. in US	2.423	2.403	2.406
Earnings C. 3 st Gen. in US	2.404	2.401	2.393

Table B.6: Counterfactual Simulation: Cost of College - Moments

Moment	Benchmark	Higher Ret. MX	Higher Ret. US
Migration Rate	0.100	0.094	0.111
College Mex. in Mex.	0.155	0.185	0.147
College 1 st Gen. in US	0.132	0.103	0.194
College 2 st Gen. in US	0.456	0.446	0.505
College 3 st Gen. in US	0.437	0.437	0.472
Earnings HS 1 st Gen. in US	1.624	1.619	1.629
Earnings HS 2 st Gen. in US	1.968	1.966	1.970
Earnings HS 3 st Gen. in US	1.963	1.960	1.967
Earnings C. Mex. in Mex.	1.222	1.238	1.219
Earnings C. 1 st Gen. in US	2.015	2.019	2.029
Earnings C. 2 st Gen. in US	2.423	2.403	2.452
Earnings C. 3 st Gen. in US	2.404	2.401	2.425

Appendix C Additional Evidence

Table C.1 reproduces the average years of schooling for each generation conditional on being high school or college educated. For Mexicans in Mexico we see that, not only they have the highest average years of education among college educated, but they also have the lowest average years of education among the high school educated. Mexicans in Mexico also have the highest returns to college, as indicated by the gap between the average earnings of college educated and high school educated. This would suggest that part of the high returns to education may be explained by heterogeneity within the groups. Using the returns to years of schooling above, and realizing that the difference in years of schooling is about 6 years larger than for second and third generation, we could say that half of the high returns can be explained by this type of heterogeneity. By not taking into account this type of heterogeneity the model and its estimation produces an upward-biased parameter measuring the skill price of college education in Mexico together with a higher cost of education. Also notice that the average years of schooling for first generation high school educated is more than three years lower than the second and third generations in corresponding educational groups. This is a possible source of bias in estimating the loss of human capital. In particular, the difference in years of education might, by itself, lead to lower earnings for high school educated first generation Mexicans compared to their second generation counterpart. In this sense, the estimated loss of human capital for this group might be biased upward. This is not true for the college group having very similar schooling years between first, second and third generations.

Table C.1: Alternative Samples II - Descriptive Statistics

	Variable	Men				Women			
		Mexico	1 st Gen	2 nd Gen	3 rd Gen	Mexico	1 st Gen	2 nd Gen	3 rd Gen
High School	Years	6.9442	7.5935	10.9782	11.1534	7.8804	7.8364	11.0349	11.3094
	s.e.	(0.0118)	(0.0326)	(0.0375)	(0.0225)	(0.0193)	(0.0517)	(0.0425)	(0.0226)
	N. obs	88803	11726	2900	6746	34084	4565	2087	5266
College	Years	16.0891	14.1232	13.7679	13.8476	15.8640	13.9926	13.8368	13.8257
	s.e.	(0.0125)	(0.0294)	(0.0227)	(0.0154)	(0.0149)	(0.0361)	(0.0228)	(0.0148)
	N. obs	13645	1871	2357	5409	10052	974	2317	5330

Standard errors in parenthesis.

Among college educated the only substantial difference is between Mexicans in Mexico and the other three generations. Among immigrants of different generations the differences

in years of schooling of college educated, although statistically significant, are very small with the larger difference being of about one third of a year between first and second generation. Table C.2 reports the returns to years of schooling, according to the results reported in this table and assuming that years of schooling have log-linear returns independently on the level of education, the difference in years of schooling between first and second generation would imply that the gap between the earnings of first and second generation with college education should be about 1% higher, from 0.4193 to about 0.43. This clearly is a very small difference. The difference between second and third generation is even smaller amounting to about 0.08 years. Again, taking into account this difference and the returns to years of schooling, the gap between between the second and third generation should be about 1% higher. The earning gaps between first and second generations mainly identify the parameters indicating the loss of human capital that first generations face, accounting for the possible bias due to heterogeneity could lead to a slighter larger gap for college educated, up to about 2%. The change would be small, and would not substantively change the results.

Table C.2 reports the regression of log hourly earnings on dummies for generations of immigrants or for Mexicans remained in Mexico, on years of schooling interacted with the previous dummies a quadratic function of experience and dummies for time and geography. The table clearly show that the returns to an extra year of schooling is higher for Mexican non-migrants than is for first generation immigrants. The returns to years of schooling for the other two groups are similar, showing a slightly flatter profile for the second compared to the third generation. The lower returns to years of schooling faced by first generation immigrants compared to all other Mexican generations rules out the possibility that the lower earnings of high school or lower educated first generation immigrants as compared to second and third generations are entirely due to lower educational attainment, in terms of completed years of education, within this group. In fact, even if the average years of schooling within the high school educated group was the same, given the lower returns to years of education we would observe a significant gap between first and second generation immigrants.

Table C.2: Returns to Years of Schooling by Generation

Dependent Var.: Log Hourly Wage				
	Non-Migrants	1 st Gen. Imm.	2 nd Gen. Imm.	3 rd Gen. Imm.
Intercept	-0.7425 (0.0168)	1.1588 (0.0278)	0.7371 (0.0635)	0.6260 (0.0497)
Years of Schooling	0.1101 (0.0005)	0.0429 (0.0012)	0.0952 (0.0049)	0.1030 (0.0035)
Experience		0.0425 (0.0016)		
Experience ²		-0.0525 (0.0028)		
N. Obs.		187914		
R ²		0.8445		

Standard errors in parenthesis.

C.1 Sensitivity of the Results to the Definition of Second and Third Generation

Table C.3 reports the results of the regression run in Table A.1 using a different data set. The data set used is taken from the Panel Study of Income Dynamics (PSID) Latino sample 1989-1990. During 1989 the PSID collected information on an additional 2043 households of Latinos to add to the original 1969 sample. Of these 2043 households about half were Mexican-Americans. Unfortunately, although it was originally planned to integrate the sample with the 1969 PSID wave and build a longitudinal study for this subgroup of individuals, the sample was later dropped and no significant longitudinal information is available for these individuals. However, this data set is the only source available on a nation-wide representative sample of Mexican-Americans which has information on the birth country of the individuals, of both their parents and all four of their grandparents. As such, this is the only data available in which an objective definition of third generation can be applied. The three columns of the table report the results of the regression using three different definition of third generation. In the first column the definition used is the widest and includes all the individuals who have at least one grandparent born in Mexico. In column 2 the definition used is stricter and requires that at least one grand parent for each parent was born in Mexico. Column 3 provides the results using the stricter definition that all grand parents were born in Mexico. While in most cases the standard errors are significantly large because of the number of individuals, the results suggest

that in general the second generation of Mexican Americans with high school or less performs similarly or better than the third generation with the same educational level. Things are a little different when we look at more educated Mexicans. In this case the returns to education seem to be more pronounced for the third generation than the second for all definitions. However, interestingly, the stricter the definition is the larger the gains from education are. It should be said that the number of observations for subgroups becomes very low for stricter definitions, and therefore the large standard errors, however the tendency is interesting in that is exactly the opposite we would expect given the criticism raised by Duncan and Trejo (2005).

Table C.3: Earnings Gaps Between First, Second and Third Generation of Mexicans - PSID Data

Dependent Var.: Log Hourly Wage	Def. 1	Def. 2	Def. 3
Other Latinos	1.9135 (0.0905)	1.9232 (0.0888)	1.9231 (0.0888)
1 st Gen. Mex. Imm.	1.6582 (0.0947)	1.6701 (0.0940)	1.6672 (0.0939)
2 nd Gen. Mexicans	1.9514 (0.1734)	1.9632 (0.1724)	1.9604 (0.1724)
3 rd Gen. Mexicans	1.9566 (0.2961)	1.9110 (0.3318)	1.5708 (0.4491)
3 rd Gen. Mex. Self Rep.	0.0826 (0.2873)	0.2975 (0.3470)	0.6235 (0.4549)
Other Latinos with Coll.	0.2503 (0.0732)	0.2272 (0.0694)	0.2298 (0.0690)
1 st Gen. Mex. Imm. Coll.	0.4103 (0.1282)	0.4090 (0.1275)	0.4091 (0.1275)
2 nd Gen. Mexicans Coll.	0.4073 (0.2390)	0.4054 (0.2378)	0.4057 (0.2378)
3 rd Gen. Mexicans Coll.	0.1297 (0.1719)	0.3827 (0.2730)	0.4288 (0.2978)
N.OBS.	518	518	518
R^2	0.9408	0.9414	0.9414

PSID Latino sample 1989-1990. Standard errors in parenthesis. The table shows the evolution of earnings across three generation of Mexicans in the US. Def.1 to Def.3 differ for how the third generation is defined. Def.1 includes all those US born individuals with US born parents and with a least one grand-parent born in Mexico; Def.2 two grand-parents; Def.3 all four grand-parents. In row 5 the dummy for third generation is interacted with a dummy for self reporting of Mexican origin. Although not significant, the estimates reported in row 5 fail to find evidence that those third generation Mexicans who identify themselves as Mexicans have lower earnings than the rest of third generation.

The last two columns of Table A.1 show that the results are robust to the definition of

second and third generations. Including in the second generation Mexicans only those who identify themselves as Mexicans - stricter definition - does not change substantially the results, as it can be seen by comparing the earnings of the second generation (rows 3 and 7) in case the standard definition is used (first two columns) or the stricter is used (last two columns). This gives some confidence that the bias induced by relying on self-identification as Mexicans to define the third generation, criticized by Duncan and Trejo (2005), leads to a bias that is negligible. Duncan and Trejo (2005) criticize the use of CPS data to infer the performance of third generation Mexican Americans because of the difficulty to objectively identify this group of individuals. They point at the fact that only those individuals that by the third generation still feel connected with their Mexican heritage answer that they are Mexicans, while those that lost this connection, for example because their Mexican parents or grandparents married White non Hispanic spouses, do not. If, as Duncan and Trejo suggest, this latter group tends to be more integrated than the former one, and perform better in the labor market, than the earnings and schooling measures reported underestimate their attainment. However, in their work Duncan and Trejo admit that they can only show a direction for the bias but cannot quantify it, while they also show that a similar bias should be present for the second generation were we not able to objectively identify it and needed to rely of the same sort of subjective information on ethnicity. The last two columns of Tables A.1 and 2 report the same regression done in the first two columns using a definition of second generation that not only relies on place of birth of parents but also on self reported ethnicity. That is, the second generation is now the subgroup of those US born individuals with at least one parent born in Mexico who identify themselves as Mexicans. In fact, as indicated by Duncan and Trejo, a small difference exists in the earnings of this subgroup. For men the coefficient in row 3 drops from 1.9694 to 1.9633, while for women drops from 1.6832 to 1.6778. For both sexes we would have a bias of about 0.6% by using the subjective definition based on ethnicity rather than the objective definition based on place of birth of parents only in the case of individuals without college. Consistently with what Duncan and Trejo sustain, increases for college educated. The coefficient in row 7 drops from 2.4240 to 2.4011, a change of about 2.3% for men and from 2.1310 to 2.1244 for women a change of about 0.7%. However, the changes are in general small and they do not affect the overall picture. In particular we still see a general worsening of the

earnings of the third generation compared to the second for college educated, although now significant only at a 10% confidence level. In sum, the last two columns of Table A.1 suggest that only a very small portion of the absence of progress, or even the regression, between the second and third generation seen in the data can be explained by the bias induced by the subjective definition of third generation based on self reported ethnicity.

Appendix D A Simplified Model

In this section I build a simplified model that, while conserving all the main characteristics of the general model presented in the paper, is analytically tractable. To build this model, I simplify the general model with two main assumptions. First, I assume that for each dynasty there are only two generations. That is, the first generation has children, but the second does not. In this case the problem becomes analytically tractable. The second assumption is that the autoregressive parameter b that governs the transmission of abilities from one generation to the other is the same for both abilities.

These assumptions clearly would alter the original model in such a way that would be impossible to derive precise answers to the questions of interest. However they do not change the main features of the original model. In particular, they maintain the fact that the decisions of an agent concerning schooling and migration are made taking into account both their current, and their children's future welfare.

D.1 The value of Migrating

I start with the problem faced by the Mexican resident who has to decide to migrate or not and his educational level,

$$\max\{v_1(s_{H,g}, s_{C,g}), v_0(s_{H,g}, s_{C,g})\}, \quad (\text{D.1})$$

where $v_1(s_{H,g}, s_{C,g})$ is the value of migrating and $v_0(s_{H,g}, s_{C,g})$ the value of staying in Mexico. The value of migrating is further decomposed in,

$$v_1(s_{H,g}, s_{C,g}) = w_1 + \beta v_{11}(s_{H,g+1}, s_{C,g+1}), \quad (\text{D.2})$$

with,

$$w_1 = \max_k \{\pi_{us,k} + s_{k,g} - z - \tau_{m,k}\},$$

Where $\tau_{m,k}$ is the cost of schooling for a Mexican immigrant. The latter component of the value of migrating is the expected value of being born in the US for the second generation, which is given by,

$$Ev_{11}(s_{H,g+1}, s_{C,g+1}) = \max_k \{\pi_{us,k} + bs_k - x + u_k - \tau_{us,k}\}. \quad (D.3)$$

Equation (D.3) takes into account the intergenerational transfer of ability given by the following law of motion

$$s_{k,g+1} = bs_{k,g} + u_k.$$

The second generation immigrant only has to choose between high school and college as being in the US for a US born person is always better than migrating to Mexico. The value of being born in the US from an immigrant father is therefore

$$Ev_{11}(s_{H,g+1}, s_{C,g+1}) = \int_{-\infty}^{\infty} \int_{-\infty}^{A+u_C} (\pi_{us,C} + bs_C - x + u_C - \tau_{us,C}) f(u_H, u_C) du_H du_C + \int_{-\infty}^{\infty} \int_{A+u_C}^{\infty} (\pi_{us,H} + bs_H - x + u_H - \tau_{us,H}) f(u_H, u_C) du_H du_C, \quad (D.4)$$

where

$$A = \pi_{us,C} - \pi_{us,H} + b(s_C - s_H) - (\tau_{us,C} - \tau_{us,H}),$$

and $f(u_H, u_C)$ is the bivariate density function of the errors. Equation (D.4) can be rearranged to obtain

$$Ev_{11}(\cdot) = \int_{-\infty}^A (A - u) f_u(u) du + \int_{-\infty}^{\infty} (\pi_{us,H} + bs_H - \tau_{us,H} + u_H) f_L(u_H) du_H - x, \quad (D.5)$$

or simply

$$Ev_{11}(\cdot) = \pi_{us,H} + bs_H - x - \tau_{us,H} + \int_{-\infty}^A (A - u) f_u(u) du, \quad (D.6)$$

where $u = u_H - u_C$. Now, equation (D.2) can be written

$$v_1(s_{H,g}, s_{C,g}) = \max\{\pi_{us,H} + s_H, \pi_{us,C} + s_C - \tau_{m,C}\} - z + \beta Ev_{11}(s_{H,g+1}, s_{C,g+1}), \quad (D.7)$$

where it is assumed $\tau_{m,H} = 0$.

D.2 The Value of Staying

Given the greater number of options the second generation have, to calculate the value of staying is a little more complicated. I start by writing

$$v_0(s_{H,g}, s_{C,g}) = w_0 + \beta Ev_0(s_{H,g+1}, s_{C,g+1}). \quad (D.8)$$

Where $w_0 = \max\{\pi_{mx,H} + s_H, \pi_{mx,C} + s_C - \tau_{mx,C}\}$. To calculate the expected value of the

second generation in this case I need to take into account the four available possibilities: migrating or not and the level of education to take. Second generations solve the following problem

$$\max_{k,I} \{\pi_{mx,H} + bs_H + u_H, \pi_{mx,C} + bs_C + u_C - \tau_{mx,C}, \pi_{us,H} + bs_H - z, \pi_{us,C} + bs_C - z - \tau_{m,C}\}. \quad (D.9)$$

The second generation, conditional on the first generation choosing to stay, choose to migrate unconditionally if and only if

$$\pi_{mx,H} + bs_H + u_H < \pi_{us,H} + bs_H - z + u_H \quad (D.10)$$

$$\pi_{mx,C} + bs_C + u_C - \tau_{mx,C} < \pi_{us,C} + bs_C - z + u_C - \tau_{m,C}, \quad (D.11)$$

or

$$\pi_{us,H} - \pi_{mx,H} > z \quad (D.12)$$

$$\pi_{us,C} - \pi_{mx,C} - (\tau_{m,C} - \tau_{mx,C}) > z. \quad (D.13)$$

However, note that if I assume that $\pi_{us,C} - \pi_{us,H} - \tau_{m,C} < \pi_{mx,C} - \pi_{mx,H} - \tau_{mx,C}$ then equation (D.13) implies equation (D.12). Let's define the left hand side of equation (D.12), $\Delta\pi_H = \pi_{us,H} - \pi_{mx,H}$, and the left hand side of equation (D.13), $\Delta\pi_C = \pi_{us,C} - \pi_{mx,C} - (\tau_{m,C} - \tau_{mx,C})$. Then I can write the expected value for a second generation Mexican of a non-migrant

$$\begin{aligned} Ev_{01}(s_{H,g+1}, s_{C,g+1}) &= \int_{-\infty}^{\Delta\pi_C} \left[E \max_k \{\pi_{us,k} + bs_k + u_k - \tau_{m,k}\} - z \right] f_z(z) dz + \\ &\int_{\Delta\pi_C}^{\Delta\pi_H} [E \max \{\pi_{us,H} + bs_H - z + u_H, \pi_{mx,C} + bs_C + u_C - \tau_{mx,C}\}] f_z(z) dz + \\ &\int_{\Delta\pi_H}^{\infty} \left[E \max_k \{\pi_{mx,k} + bs_k + u_k - \tau_{mx,k}\} \right] f_z(z) dz, \quad (D.14) \end{aligned}$$

setting

$$\begin{aligned} A_1 &= \pi_{us,C} - \pi_{us,H} + b(s_C - s_H) - \tau_{m,C}, \\ A_z &= \pi_{mx,C} - \pi_{us,H} + b(s_C - s_H) + z - \tau_{mx,C}, \end{aligned}$$

and,

$$A_2 = \pi_{mx,C} - \pi_{mx,H} + b(s_C - s_H) - \tau_{mx,C}.$$

After some algebra

$$Ev_{01}(\cdot) = \pi_{us,H} + bs_H + F_z(\Delta\pi_C) \left[\int_{-\infty}^{A_1} (A_1 - u) f_u(u) du \right] - \int_{-\infty}^{\Delta\pi_C} z f_z(z) dz +$$

The value of not migrating for the father is

$$\int_{\Delta\pi_C}^{\Delta\pi_H} \left[\int_{-\infty}^{A_z} (A_z - u) f_u(u) du \right] f_z(z) dz + [1 - F_z(\Delta\pi_H)] \left[\int_{-\infty}^{A_2} (A_2 - u) f_u(u) du \right],$$

$$v_0(\cdot) = \max_k \{ \pi_{mx,k} + s_k - \tau_{mx,k} \} + \beta E v_{01}(\cdot). \quad (\text{D.15})$$

D.3 The Selection Mechanism

In order to show that the model is identified by the available data, I need to show the relationship between the incentive to migrate and its cost together with the relationship between the incentive to migrate and the abilities. I need to verify that in all possible situations an increased cost of migrating decreases the incentive to migrate, and check in which direction the selection goes in terms of abilities and how the selection process reacts to a change in the cost of migration.

First note that a person migrates if and only if

$$v_1(\cdot) - v_0(\cdot) > 0, \quad (\text{D.16})$$

where the value of moving is given by

$$v_1(\cdot) = \max_k \{ \pi_{us,k} + s_k - \tau_{mx,k} \} - z + \beta E v_{11}. \quad (\text{D.17})$$

Therefore, a higher z implies a lower probability to migrate. Another important feature to study is how the incentive to migrate changes with both abilities. Taking the derivative of $v_1 - v_0$

$$\frac{\partial(v_1 - v_0)}{\partial s_H} = \frac{\partial(w_1 - w_0)}{\partial s_H} + \beta \frac{\partial(E v_{11} - E v_{01})}{\partial s_H}. \quad (\text{D.18})$$

Let's first concentrate on the second part.

$$\frac{\partial(E v_{11} - E v_{01})}{\partial s_H} = b[1 - F_u(A)] - \frac{\partial E v_{01}}{\partial s_H}, \quad (\text{D.19})$$

where

$$\frac{\partial E v_{01}}{\partial s_H} = b \left\{ 1 - F_u(A_1) F_z(\Delta\pi_C) + \int_{\Delta\pi_C}^{\Delta\pi_H} F_u(A_z) f_z(z) dz - [1 - F_z(\Delta\pi_H)] F(A_2) \right\}. \quad (\text{D.20})$$

Rearranging the terms it is possible to write

$$\frac{\partial(Ev_{11} - Ev_{01})}{\partial s_H} = b \left\{ [1 - F_z(\Delta\pi_H)][F(A_2) - F(A_1)] + \int_{\Delta\pi_C}^{\Delta\pi_H} [F_u(A_z) - F(A_1)]f_z(z)dz - [F(A) - F(A_1)] \right\}. \quad (D.21)$$

Assuming that $\pi_{us,C} - \tau_{us,C} - (\pi_{us,H} - \tau_{us,H}) > \pi_{mx,C} - \tau_{mx,C} - \pi_{mx,H}$ it is possible to prove that the derivative with respect to the manual ability is always negative, while the intellectual ability increases the incentive to migrate. In fact, the derivative with respect to the manual ability s_H is

$$\frac{\partial(Ev_{11} - Ev_{01})}{\partial s_H} = b \left\{ [1 - F_z(\Delta\pi_H)][F(A_2) - F(A_1)] + \int_{\Delta\pi_C}^{\Delta\pi_H} [F_u(A_z) - F(A_1)]f_z(z)dz - [F(A) - F(A_1)] \right\}. \quad (D.22)$$

First note that $z = \Delta\pi_C$ implies $A_z = A_1 + \tau_{m,C} - \tau_{mx,C}$. Therefore,

$$\int_{\Delta\pi_C}^{\Delta\pi_H} F(A_z)f_z(z)dz \leq [F_z(\Delta\pi_C) - F_z(\Delta\pi_H)]F(A_1), \quad (D.23)$$

and

$$\int_{\Delta\pi_C}^{\Delta\pi_H} [F_u(A_z) - F(A_1)]f_z(z)dz \leq 0. \quad (D.24)$$

Thus,

$$\frac{\partial(Ev_{11} - Ev_{01})}{\partial s_H} \leq b \{ [1 - F_z(\Delta\pi_C)][F(A_2) - F(A_1)] - [F(A) - F(A_1)] \}. \quad (D.25)$$

The rhs of equation (D.25) is clearly negative if $A > A_2$. In other words, if the difference in the spread of earnings is lower than the difference in the spread of costs, increasing the lower ability makes the expected value of remaining in Mexico greater than going to the US. Intuitively, if the first generation decides to migrate, the second generation pays the cost of education in the host country. As such, the value of being born in the US for the second generation is lower the greater is the intellectual ability and the lower is the manual ability. This arises because the cost for a college education is relatively lower compared to its gain than for high school.

We are now left with the first part of the derivative,

$$\frac{\partial(w_1 - w_0)}{\partial s_H} \quad (\text{D.26})$$

In this case we cannot find a unique solution in terms of its sign. If the educational choice of the individual is the same no matter the location she chooses, then the derivative in equation (D.26) is zero. For example, if

$$s_H - s_C > \pi_{mx,C} - \pi_{mx,H} - \tau_{mx,C} > \pi_{us,C} - \pi_{us,H} - \tau_{m,C}$$

then the individual endowed with these two levels will choose to stay high school educated no matter where will end up living. Or, alternatively, if

$$\pi_{mx,C} - \pi_{mx,H} - \tau_{mx,C} > \pi_{us,C} - \pi_{us,H} - \tau_{m,C} > s_H - s_C$$

then the individual will choose to acquire a college education again independently of migrating or not. In both these cases a small change of s_H that does not alter the inequality does not change the difference in values of migrating or not migrating due to the part relative to the gain of the current generation. A third, intermediate, case is if

$$\pi_{mx,C} - \pi_{mx,H} - \tau_{mx,C} > s_H - s_C > \pi_{us,C} - \pi_{us,H} - \tau_{m,C}.$$

In this case an individual would choose to acquire a college education only if the z is high enough to prevent her to migrate, otherwise being high school educated is better. The difference in values of migrating or not migrating due to the gain of the current generation is now positive when s_H increases.

If we take the derivative with respect to the intellectual ability, we obtain the same but reversed results. Conditional on choosing college education in the host country, the selection is necessarily positive. In fact, an individual who chooses to be college educated and migrate would choose to be college educated even if the cost of migrating is too high to make it worthwhile. In this case, a small change in both abilities does not change the gain of migration to the current generation. However, the gain of future generations is increasing in the intellectual ability of the current generation, which determines positive selection. Overall, the derivatives of migration gain with respect to both abilities are in general undetermined. When the school choice does not change with the migration decision the derivative with respect to s_H (s_C) is negative (positive), but it can be positive (negative) when, upon migration, the education decision is changed to high school rather than college. However, provided that the population

of switchers is relatively small compared to the rest of the population^{D.1}, the model is clearly identified by the data moments.

Figure D.1: Migration Choice Based on Abilities

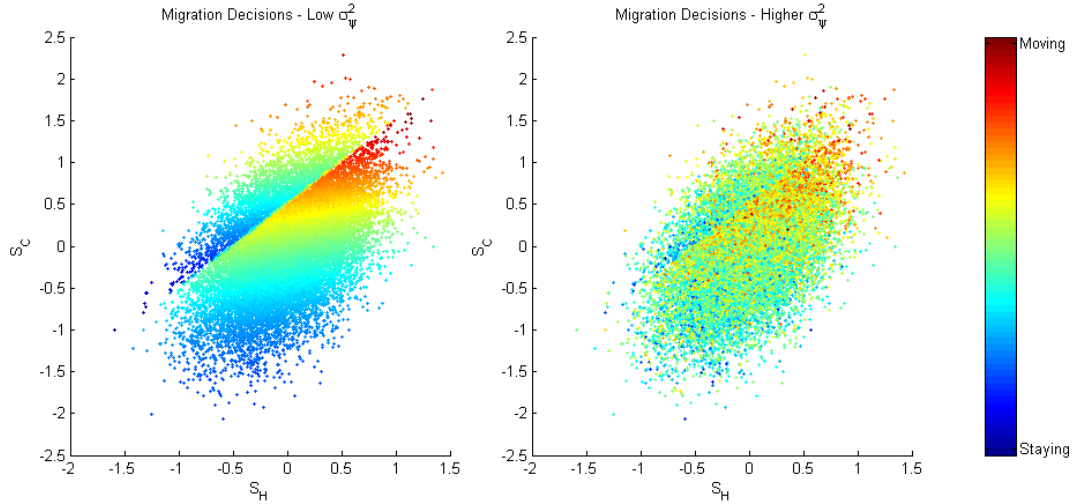


Figure D.1 shows, for different values of abilities, the gain from migration. The figure represents a simulation using the model with one z and one x with the estimated parameters except for the variance and the disutility cost of migration. In the color scale warmer colors indicate higher values for the gain of migration. The first panel simulates the model under a very low variance of the disutility shock distribution, while in the second panel the variance is larger. In both panels we can clearly distinguish two regions separated by positively sloped line. The two regions contain individuals making different choices about their education. In the bottom right there are individuals choosing to remain with a high school education, while in the up left region individuals are college educated. Figure D.2 given more details on the schooling decisions. From the first panel we can clearly see the selection pattern, conditional on schooling, low levels of intellectual ability lead to lower gains from migration, while higher levels make people hotter about moving. The opposite is true for the manual ability, although the selection is very mild. The right panel shows the same pattern, but the selection is more blurred due to the higher variance. The figure clearly shows that a lower variance makes the

^{D.1}Figure D.2 shows the distribution of individuals according to their schooling choices conditional on migrating or non migrating. The figure clearly shows that the population of switchers is very small compared to the other group.

selection mechanism stronger.

Figure D.2: Schooling and Migration Choices Based on Abilities

