INTERNATIONAL ECONOMIC REVIEW
Vol. 52, No. 2, May 2011

# INTERGENERATIONAL TRANSMISSION OF ABILITIES AND SELF-SELECTION OF MEXICAN IMMIGRANTS*

BY VINCENZO CAPONI[1]

*Ryerson University, Canada*

This article 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 sacrifices made by the first generation of Mexican immigrants for the benefits of future generations. In particular, the estimation results suggest that there is a significant one time loss of human capital of immigrants upon migration, that migrants are positively selected from the ability distribution, and that they transmit substantial human capital to their children. Finally, the model is used to evaluate the effects of social policies designed to reduce the "brain drain" from Mexico to the United States.

## 1. INTRODUCTION

This article seeks to better understand the performance of Mexican immigrants in the United States and their descendants. If we focus soley on immigrants, the following three facts can be observed.[2] First, the earnings of Mexican immigrants in the United States are higher than the earnings of nonmigrant Mexicans. Earnings continue to increase from the first to the second generation of Mexicans in the United States, but then stop increasing or regress from the second to the third generation.[3] Second, conditional on education, the earnings differential is higher for high school educated than for college educated immigrants.[4] Average 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 whereas college earnings decrease. Third, if we measure educational attainment in terms of the share of college educated individuals, Mexican immigrants are less educated than nonmigrant Mexicans. Second generation Mexicans living in the United States substantially improve their education compared to their parents, whereas attainment for the third generation is slightly lower than for the second.

This article aims to explain this evidence using a structurally estimated intergenerational self-selection model of migration and education. The estimation results indicate that it is important

* Manuscript received September 2007; revised June 2009.

[1] 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 article. 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. Please address correspondence to: Vincenzo Caponi, Department of Economics-Faculty of Arts, Reyerson University, 350 Victoria street, Toronto, ON, M5B 2K3, Canada. Phone: 417 979 500 (ext 6187). fax: 416 598 5916. Email:vcaponi@ryerson.ca

[2] The data used are from the Current Population Survey (CPS)—March supplement. The survey collects information on foreign-born individuals independently of their legal status, including information on legal and illegal immigrants. Chiquiar and Hanson (2005) show that accounting for the possible undercount of illegal immigrants in the 2000 Census data does not change significantly the statistics of immigrants in general.

[3] First generation immigrants are all Mexican-born individuals who migrated to the United States at the age of 22 or older. Second generation are all U.S.-born individuals with at least one Mexican-born parent. Third generation are 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.

[4] The definition I use of high school education includes all individuals with a high school degree or less; college educated are those who attended a college for at least one term.

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 and contributes to their future earnings. The estimated model shows that the first function is generally negatively affected by migration, whereas the second is not. That is, immigrants face a significant loss of earnings capacity upon migration, but there is no loss of capacity to transfer human capital to their children. For example, Mexican immigrants may have difficulty adapting to the U.S. labor market because of the different language or different cultural traits. However, human capital is also based on intelligence, work ethic, and physical or other attributes that can be easily transferred from parents to children regardless to the migrant status. This distinction explains the large gap in earnings between the 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 greatly affected and is greater than the earnings of the first generation would suggest. Moreover, second-generation Mexicans do not face the same difficulty the first generation faces; therefore they are capable of fully utilizing the inherited human capital, significantly improving their performance compared to their parents. The estimation results also indicate that the loss of human capital is higher for college-educated immigrants than for immigrants with a high school education or less, which explains why the returns from migration for college educated are lower.

Furthermore, this article 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. The hypothesis proposed by Borjas that immigrants are negatively selected is inconsistent with the observations on the second and third generations 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 article suggests that because of positive self-selection, human capital within the second generation is higher than the first on average, and so are the earnings, whereas the third and successive generations show a reversion to the mean.

Even among the high school educated, I find that immigrants are positively selected. Given the difficulty in adapting their own human capital to the U.S. labor market, particularly for college-educated immigrants, it may be optimal for them to avoid 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 United States. However, because of altruism toward future generations, migrating is an optimal choice for them due to the benefits for 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 nonmigrants.

Finally, the article evaluates alternative policies aimed at integrating immigrants into the U.S. labor market that may have positive impacts on the human capital distribution in the host country by mitigating the loss that immigrants face upon migration or aimed at mitigating or reducing the brain drain from Mexico to the United States. I propose two scenarios that could result from policies aimed at integrating immigrants faster, one in which high school educated are targeted and the other in which are targeted 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 generations by strengthening the positive self-selection process of immigrants. I also propose one policy that reduces the brain drain by reducing the cost of education in Mexico. The simulation results indicate that this policy is effective. Immigrants become less self-selected with respect to human capital, and among nonmigrant Mexicans the average human capital increases. Moreover, the policy also reduces the migration rate, especially among the highly educated Mexicans.

The article 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 article related to Mexican generations of immigrants in the United States 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 who 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.[5] 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 postsecondary education than the first and third generations. 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 generations.

One contribution of my article is to the literature on the selection mechanism of immigrants in a destination country. Borjas (1987, 2000) found that recent cohorts of immigrants in the United States are negatively selected. Among them are Mexicans, who represent about 30% of the whole foreign-born population in the United States. Chiquiar and Hanson (2005), on the contrary, suggest that Mexican immigrants are positively or intermediately 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.[6] Moreover, consistent 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 find strong support that this being 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 United States and compares the performance of these generations with

[5] "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" says Card (2005 footnote 1, p. 2).

[6] The level of significance of the relevant parameters is much higher for those of 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.

TABLE 1
SAMPLE—DESCRIPTIVE STATISTICS

| | Men | | | | Women | | | |
|---|---|---|---|---|---|---|---|---|
| Variable | Mexico | 1st Gen | 2nd Gen | 3rd Gen | Mexico | 1st Gen | 2nd Gen | 3rd Gen |
| Log wages | 0.6831 | 2.1796 | 2.5447 | 2.5641 | 0.7585 | 1.9768 | 2.3613 | 2.3550 |
| | (0.0024) | (0.0046) | (0.0082) | (0.0053) | (0.0038) | (0.0064) | (0.0083) | (0.0054) |
| College share | 0.1539 | 0.1337 | 0.4541 | 0.4369 | 0.2356 | 0.1768 | 0.5265 | 0.4970 |
| | (0.0011) | (0.0029) | (0.0069) | (0.0045) | (0.0020) | (0.0051) | (0.0075) | (0.0049) |
| Age | 36.5823 | 39.8096 | 36.7312 | 38.0781 | 34.9013 | 41.7280 | 36.7970 | 38.5028 |
| | (0.0349) | (0.0898) | (0.1629) | (0.0954) | (0.0470) | (0.1327) | (0.1747) | (0.0997) |
| School years | 8.3826 | 8.4668 | 12.2449 | 12.3304 | 9.8088 | 8.9248 | 12.5101 | 12.5600 |
| | (0.0147) | (0.0342) | (0.0299) | (0.0187) | (0.0224) | (0.0534) | (0.0315) | (0.0182) |
| Population | 11,031,430 | 1,229,532 | 500,392 | 1,086,856 | 5,226,140 | 441,129 | 388,572 | 895,240 |
| *N* obs | 102,448 | 13,597 | 5,257 | 12,155 | 44,136 | 5,539 | 4,404 | 10,596 |

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

Mexican nonmigrants—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 United States. The table presents average log hourly earnings and educational attainment 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, for Mexicans immigrants in the United States.[7]

Table 1 shows that there is a large difference between the log earnings of Mexicans working in Mexico and Mexicans in the United States. The first generation of male immigrants earns 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 United States, with the second and third generations of males earnings about 0.37 log-points more. Table 1 also shows that the first generation immigrants have a much lower level of education than their descendants. Although 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%. Earnings from the second to the third generation increase slightly by about 2%.[8]

To explore the relationship between education and the intergenerational evolution of the earnings of immigrants, I estimate a regression of log hourly earnings on a set of dummy 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] Table 2 reports the earnings differentials associated with migration and education as well as the evolutions of earnings across all generations.[10] The table reports results for men and women using two definitions: the standard definition and a stricter definition.

[7] See the Appendix for details on the sample selection.

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

[9] The introduction of the year of birth in the regression allows us to net out cohort and business cycle effects on earnings. The year of birth can be used for CPS data since there are 15 pooled surveys from 1994 to 2008, but not for Mexican data, whereas the inclusion of a polynomial function of age centered at 48 allows us to see the results as a proxy for lifetime earnings (see Altonji and Doraszelski, 2005, for details).

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

TABLE 2
INTERGENERATIONAL EVOLUTION OF EARNINGS

| | Standard Definition | | Stricter Definition | |
|---|---|---|---|---|
| Dependent Var.: Log Hourly Wage | Men | Women | Men | Women |
| Returns to college | | | | |
| 1st gen. imm. | 0.3832 | 0.3654 | 0.4259 | 0.3923 |
| | (0.0232) | (0.0341) | (0.0197) | (0.0291) |
| 2nd gen. Mexicans | 0.4546 | 0.4510 | 0.4389 | 0.4505 |
| | (0.0284) | (0.0324) | (0.0295) | (0.0341) |
| 3rd gen. Mexicans | 0.3995 | 0.4504 | 0.3992 | 0.4511 |
| | (0.0175) | (0.0193) | (0.0171) | (0.0189) |
| Mexicans in Mexico | 1.2230 | 1.1440 | 1.2232 | 1.1440 |
| | (0.0081) | (0.0118) | (0.0079) | (0.0116) |
| Returns to migration | | | | |
| without college | 1.6263 | 1.3507 | 1.6231 | 1.3629 |
| | (0.0194) | (0.0274) | (0.0167) | (0.0249) |
| With college | 0.7865 | 0.5721 | 0.8259 | 0.6112 |
| | (0.0284) | (0.0401) | (0.0245) | (0.0354) |
| Intergenerational dynamics | | | | |
| From 1st to 2nd gen. no coll. | 0.3479 | 0.3194 | 0.3445 | 0.3140 |
| | (0.0212) | (0.0270) | (0.0210) | (0.0267) |
| From 1st to 2nd gen. with coll. | 0.4193 | 0.4050 | 0.3575 | 0.3723 |
| | (0.0301) | (0.0386) | (0.0288) | (0.0362) |
| From 2nd to 3rd gen. no coll. | −0.0100 | −0.0317 | −0.0040 | −0.0265 |
| | (0.0226) | (0.0265) | (0.0227) | (0.0268) |
| From 2nd to 3rd gen. with coll. | −0.0652 | −0.0323 | −0.0437 | −0.0260 |
| | (0.0246) | (0.0269) | (0.0256) | (0.0283) |

The standard definition of the second generation is the one provided above, which identifies the second generation as all those U.S.-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.

The first four rows present the earnings differential between college and high school educated Mexicans for each generation. Focusing on the first column, it can be noted that the earnings differentials are quite high for each generation of men, but are much higher for Mexican nonmigrants. Among the first generation of Mexicans in the United States, first row, those men with college have an average earnings 0.38 log-points larger than those with only high school or less. The differential 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 earnings differential between college and high school educated in Mexico.[11]

Rows 5 and 6 present the earnings differential between migrants and nonmigrant Mexicans calculated by looking at the differences between the log earnings of Mexicans living in Mexico and the first generation Mexicans in the United States for each schooling category. Row 5 shows that immigrants with only high school in the United States earn on average 1.63 log-points more than Mexicans living in Mexico, whereas for college educated the gain is only 0.79. That is, the earnings differential between migrant and nonmigrant Mexicans is much lower for college educated than for high school educated Mexicans. Rows 7–10 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

[11] 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. 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.

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.[12]

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 earnings differential between college and high school educated Mexicans for the second generation in row 2, columns 1 and 3 and the 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. The facts shown in this section motivate the next section of the article, which introduces a theoretical model capable of replicating the features of the data.

## 4. MODEL

This section presents 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 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.[13] 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 United States, as observed in the data.[14]

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 United States, and (4) college and working in the United States. 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, whereas 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 to be permanent and uncorrelated with abilities. The ability cost makes part of the ability endowments of immigrants unusable for producing earnings in

[12] 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). For more evidence of this heterogeneity and a discussion of its implications for the interpretation of the results see the working paper version of this article, Caponi (2009, Tables C.1 and C.2 in Appendix C).

[13] See Heckman et al. (2006) for the role of cognitive and noncognitive 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.

[14] This 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 would have an incentive to migrate to the United States and not vice versa.

the host country. The psychic cost is known at the time the decision to migrate is made, as are the returns; therefore the model does not allow for return migration.[15]

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.[16] 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.[17] 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.[18]

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 his or her abilities, the psychic cost of 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 $I_g$ is an indicator that takes the 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

[15] Not considering return migration can induce bias in the estimation of some of the parameters if the migrants that return to Mexico are also self-selected. This represents a problem for the estimation results. However, the main results, particularly the identification of the selection mechanism, are identified by the differences between second and third generations of Mexicans, which should not be affected by return migration.

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

[17] The 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.

[18] Behrman and Rosenzweig (2002) find that the positive relationship between parents and children's schooling, documented elsewhere in the literature, disappears, or becomes negative, for mothers when appropriately controlling for unobserved components. The relationship remains positive for fathers, but the causal effect is small. These findings support the assumption that are primarily ability endowments to be transferred across generations.

each generation are represented by

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

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:

$$s_{k,g+1} = b_k(s_{k,g} - \xi z I_g) + u_{k,g+1}, \tag{2}$$

were $\xi$ 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:

$$\begin{pmatrix} u_H \\ u_C \end{pmatrix} \sim N \begin{pmatrix} 0\,| & \sigma_H^2 & \sigma_{HC} \\ 0\,| & \sigma_{HC} & \sigma_C^2 \end{pmatrix}.$$

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

$$w_{mx,C,g} - \tau_{mx,C} > w_{mx,H,g}, \tag{3}$$

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 forgone 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

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

In contrast, immigrants choose to attend college if

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

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

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

where $\tau_{us}$ is the cost of a college education for a U.S.-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 United States; and the cost for a U.S.-born individual, either of the second or third generation. Because the cost of education reflects forgone 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 nonmigrants and the second and third generations of Mexicans in the United States. The direct and psychic costs faced by immigrants may be close to the ones faced by nonmigrants given that the education is acquired in Mexico, whereas forgone earnings are comparable to the ones faced by second and third generations given that they work in the United States. 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 United States. Assuming log utility and given that one period is equivalent to one generation, the value is given by

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

where $v_{us}$ is the value of being born in the United States, $\beta$ is the parameter that measures altruism, and $\psi$ is a utility cost of migrating drawn at the time the migration decision is made from a normal distribution with mean $\mu_\psi$ and variance $\sigma^2_\psi$. 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

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

The value $(v_{us})$ of being born in the United States is given by

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

An agent who 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:

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

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

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

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

The model is estimated by simulated method of moments McFadden (1989). Mayer (2008) estimates the same model without migration using data on Americans, and I assume that the process that determines the intergenerational transmission of abilities is identical for Americans and for Mexicans. This allows me to use some of the parameters estimated by Mayer (2008) and to concentrate on estimating the remaining parameters related to the behavior of Mexicans. The estimation performed here can therefore 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 extended to the simulated method of moments by Gourinchas and Parker (2002).

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 $\chi$. 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

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

where $W_\chi^{-1}$ 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 nonmigrant Mexicans as well as first-, second-, and third-generation Mexican immigrants in the United States. The set of parameters can also be partitioned into two sets, $\theta$ and $\gamma$, such that the set of parameters $\theta$ 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 $\theta$, it is possible to estimate $\gamma$ independently and use the estimates in a second stage to estimate $\theta$. The parameters Mayer estimates, taken as coming from the first stage, are

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

whereas I estimate the following set of parameters

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

Note that the skill prices in the United States 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 United States and Mexico, and I do not use data on Americans. Although I do assume that the intergenerational transmission process is the same for Americans and Mexicans, it is not necessary to assume that Americans and Mexicans in the United States face the same set of skill prices. Neighborhood effects, discrimination, and other factors could make a Mexican with the same abilities as an American face a different wage. Therefore, in what follows, $\pi_{us,k}$ should be interpreted as the skill prices for Mexicans in the United States.[19]

There are two 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 $\beta$ and the skill price of high school ability in Mexico $\pi_{mx,H}$. Because 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. Assuming that a period is about

[19] In the first stage, Mayer estimates the difference between the skill prices for Americans. Footnote 33 shows that the estimated difference of the skill prices for Mexicans in the United States is not significantly different from the estimated difference of the skill prices for natives. Therefore, it cannot be excluded that Mexicans and natives in the United States face the same set of prices, or at least the same returns from a college education.

30 years long, the value reflects a discount factor of 0.9615 per year, which would generate an interest rate equal to 0.04.[20,21]

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

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

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, 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 $\theta$ be $G_\theta$ and the Jacobian of the same moment functions with respect to $\gamma$ be $G_\gamma$. Let $\Omega_\gamma$ be the covariance matrix of the $\gamma$ estimates and $\Omega_g$ the covariance matrix of the data moments. It can be proved[22] that a consistent estimator of the covariance matrix of $\theta$ in a 2SSMM procedure is obtained by

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

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:

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

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 $\Omega_\gamma$ of the first step estimates increases, and also increases if $G_\gamma$, the sensitivity of the second stage moments to the first stage estimates, is higher.

5.1. *Identification Strategy.* Mayer (2008) provides a discussion of the identification of the parameters $\gamma$ in the first stage estimation. He uses data from the Panel Study of Income Dynamics (PSID) that make it 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. Among the main moments he uses in his SMM estimation procedure are the correlation between earnings and school choices of parents and children. These moments cannot be computed with the data I use since they do not identify parent–child couples. As Mayer shows, these correlations are essential to identify all of the parameters of the ability distribution $\sigma_H$, $\sigma_C$, $b_H$ and $b_C$, and $\rho$. For this reason I rely on the PSID data and on the work already done by Mayer.[23]

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

[20] Notice 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.

[21] In Section 6, I conduct a sensitivity analysis on the model with alternative values of $\beta$, which shows that, although 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.

[22] See Laibson et al. (2005) for a proof based on Newey and McFadden (1994) and Gourinchas and Parker (2002).

[23] Alternatively I could have included the same data in my own estimation performing a one-step estimation. However, the one-step estimation would have been computationally quite intensive.

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 1st gen. in United States | 0.1337 | 0.0009 |
| College 2nd gen. in United States | 0.4541 | 0.0013 |
| College 3rd gen. in United States | 0.4369 | 0.0013 |
| Earnings HS 1st gen. in United States | 1.6263 | 0.0103 |
| Earnings HS 2nd gen. in United States | 1.9742 | 0.0193 |
| Earnings HS 3rd gen. in United States | 1.9642 | 0.0129 |
| Earnings C. Mex. in Mex. | 1.2230 | 0.0099 |
| Earnings C. 1st gen. in United States | 2.0095 | 0.0238 |
| Earnings C. 2nd gen. in United States | 2.4288 | 0.0210 |
| Earnings C. 3rd gen. in United States | 2.3636 | 0.0140 |

NOTE: See the Appendix for a detailed explanation of how the moments are calculated.

in the table is the same used for Table 2.[24] 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 from the second to the fifth row 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 nonmigrant Mexicans with a high school education or lower. The second column of Table 3 shows the standard error for each moment.

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 determine it, is based on the differences between the third- and the second-generation moments. Because all the parameters faced by the third and the second generations are the same, the differences between the second and the third generation moments are due to the changes in the ability distribution only.[25] This, together with the initial conditions on the ability distribution for the generation of Mexicans nonmigrants 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 nonmigrant Mexicans, that 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 United States 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. Moreover, the knowledge of the distribution of abilities across generations allows us to identify the parameters from the moments on earnings and educational attainment. From the earnings of the second generation we can identify the parameter describing the loss of ability to transmit

[24] See the Appendix for a detailed description of the sample used and how the moments and the covariance matrix are derived.

[25] The model assumes a stationary intergenerational environment. Increased returns to education, lower costs of communication and transportation, and changes in immigration policies may have affected migration decisions of different generations of migrants. However, as documented by Kurokawa (2006) and Binelli (2009), the skill premium increased similarly in both the United States and Mexico, keeping the relative returns similar over time. The reduction of migration costs is negligible in terms of lifetime earnings in the United States. The abolition of the Bracero program in 1964, which likely increased the cost of migrating to the United States disproportionately to low skill Mexicans, also sparked illegal immigration. Although the change in policy changed the status of newer Mexican immigrants, it did cause the same change to their number and type.

TABLE 4
FIRST STAGE PARAMETER ESTIMATES

| Parameters | Point Estimate | S.E. |
|---|---|---|
| $b_H$ | 0.1118 | 0.0669 |
| $b_C$ | 0.5249 | 0.0417 |
| $\rho$ | 0.5154 | 0.1318 |
| $\sigma_H$ | 0.3776 | 0.0135 |
| $\sigma_C$ | 0.5624 | 0.0346 |

NOTE: Mayer (2008, Table 2).

human capital $\xi$, whereas the cost of a college education in the United States is identified by the share of college educated. From the earnings of the first generation we can identify the loss of human capital $z$, whereas from the share of the college educated we can identify the cost of college for the first generation. From the moments of Mexicans nonmigrants we can identify 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.

## 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 ($b_H = .11$).

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.[26] 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 is 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.[27] 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, Model 3 is a restricted version of Model 2, where the $\xi$ 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.[28,29] 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. Indeed, a

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

[27] Standard 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.

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

[29] A formal overidentifying restriction test can be done on the models by testing the null hypothesis $\eta(\hat{\theta}, \hat{\gamma}) = 0$; all the models are rejected. A possible reason for the rejections is that the model is not flexible enough to capture the differences between the two types of occupations, especially among immigrants.

TABLE 5
PARAMETER ESTIMATES

| | Model 1 | | | Model 2 | | | Model 3 | |
|---|---|---|---|---|---|---|---|---|
| Parameter | 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 |
| $\tau_{mx}$ | 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 |
| $\tau_m$ | 0.554 | 0.113 | | 0.258 | 0.333 | | 0.260 | 0.122 |
| $\tau_{us}$ | 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 |
| $\xi$ | 0.006 | 0.227 | | 0.003 | 0.063 | | | |
| $\mu_\psi$ | 2.920 | 0.104 | | 2.927 | 0.020 | | 2.925 | 0.058 |
| $\sigma^2_\psi$ | 0.359 | 0.075 | | 0.323 | 0.013 | | 0.323 | 0.023 |
| $\eta(\hat{\theta}, \hat{\gamma})$ | | 147.88 | | | 9.43 | | | 11.86 |

formal test based on the objective function rejects the hypothesis of one $z$ against two $z$'s.[30] Finally, the value of $\xi$ estimated in Model 2 is close to zero with a relatively small standard error. This finding indicates that there is no loss of capacity to transmit human capital from the first to the second generation of Mexicans and that the loss faced by the first generation is transitory. In fact, if the loss were permanent, then the parameter $\xi$ would have had an estimated value close to 1, but such a high value is clearly rejected by the estimation in Model 2. 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 parameter 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 ($\tau_{mx}$) is about 1.09. Assuming that the present value of the lifetime earnings of a high school or less educated nonmigrant is about 50,000 U.S. dollars,[31] the return to college is about 41,000 U.S. dollars, implying an average cost of attending college of about 60,000 U.S. dollars.[32] 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 U.S. dollars contribute to explaining the relatively 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 may 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.

[30] The difference between the objective function of Models 1 and 2 is a chi-square statistic with one degree of freedom. Its value is 138.45, which rejects the restriction in Model 1.

[31] The figure is obtained using the value of the intercept for men in Table A.1 in the Appendix, together with the coefficient estimates of the centered polynomial on age (age-48): $earnings(age) = 0.6474 - 0.05 * (age - 48) - 0.0005* (age - 48)^2$. Earnings are calculated assuming a high school or lower educated Mexican resident works for 43 and discounting using an yearly interest rate of 3%. The figure is also adjusted to reproduce a value in U.S. dollars PPP adjusted.

[32] The 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.

For the same reasons, immigrants face higher educational costs than Americans. Their forgone earnings are closer to the forgone earnings of Americans since they have the option to migrate earlier and work in the United States 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 nonmigrants face.

Rows 3 and 4 show the skill prices for Mexicans working in the United States. The skill price for high school educated Mexicans working in the United States ($\pi_{us,H}$) is 1.94, whereas for college educated Mexicans in the United States ($\pi_{us,C}$) it is 1.96, both significantly different from zero at the 5% level. The difference between the college and the high school skill price for Mexicans born in the United States measures the return to college net of self-selection, which in this case is about 2% and not significantly different from zero.[33] Row 5 in the table shows the cost of college attendance for Mexican immigrants ($\tau_m$), which is estimated at 0.26, equivalent to about 81,000 U.S. dollars, whereas Row 6 shows the cost for U.S.-born Mexicans ($\tau_{us}$), estimated at 0.11, equivalent to 37,000 U.S. dollars.

Rows 7–10 present the estimates of the loss of human capital. Row 7 shows the direct loss faced by immigrants with a high school education ($z_H$) whereas Row 8 shows this for a college educated immigrant ($z_C$); they are, respectively, 0.35 and 0.64. These estimates are both significantly different from zero at the 5% level, and they are also significantly different at the same confidence level.[34] Therefore, the model suggests that the intellectual ability is more difficult to adapt to the U.S. 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 ($\xi$) 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.[35]

As mentioned in Section 5.1 the parameter $\beta$ is not estimated; rather the value of 0.3079 is assigned to it. To evaluate the importance of this parameter I perform a sensitivity analysis and reestimate the model with two alternative values: 0.154 – Model Beta-1 and 0 – Model Beta-2. Table 6 shows the reestimated 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 $\beta$ 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.

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.[36] 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

[33] Although 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 United States and Americans are not significantly different.

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

[35] Other than the parameters $z$ and $\xi$, 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.

[36] See the working paper version of this article, Caponi (2009, Table B.2 in Appendix B) for 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 able to reproduce: the college share and the earnings of the second generation.

TABLE 6
PARAMETER ESTIMATES

| | Benchmark Model | | Model Beta-1 | | Model Beta-2 | |
|---|---|---|---|---|---|---|
| Parameter | 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 |
| $\tau_{mx}$ | 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 |
| $\tau_m$ | 0.260 | 0.122 | 0.253 | 0.461 | 0.276 | 0.165 |
| $\tau_{us}$ | 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 |
| $\mu_\psi$ | 2.925 | 0.058 | 2.148 | 0.072 | 1.674 | 0.013 |
| $\sigma^2_\psi$ | 0.323 | 0.023 | 0.140 | 0.050 | 0.038 | 0.012 |
| $\eta(\hat{\theta}, \hat{\gamma})$ | | 11.86 | | 13.69 | | 22.28 |

TABLE 7
AVERAGE ABILITIES AND SELF-SELECTION

| Decision | Manual Ability (H) | Intellectual Ability (C) |
|---|---|---|
| Migrant | 0.025 | 0.049 |
| Nonmigrant | −0.003 | −0.009 |
| Migrant with high school | 0.042 | −0.051 |
| Migrant with college | −0.088 | 0.711 |
| Nonmigrant with high school | 0.016 | −0.128 |
| Nonmigrant with college | −0.109 | 0.641 |

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 $\beta$ and $\sigma^2_\psi$ in the same direction, 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.

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 thereby evaluate the amount of human capital brought to the United States 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 nonmigrants in Mexico unconditional and conditional on the educational choice.

Rows 3 and 5 of Table 7 show the averages of the manual and intellectual ability conditional on high school for migrants and nonmigrants, respectively. For the manual ability, the average ability for migrants is higher than for nonmigrants, 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

TABLE 8
HUMAN CAPITAL ACCOUNTING

| | Benchmark | | | Counterfactual $\xi$ | | |
|---|---|---|---|---|---|---|
| Generation | (H) | (C) | (A) | (H) | (C) | (A) |
| Nonmigrant | 50.00 | 172.85 | 68.99 | 49.97 | 173.30 | 69.51 |
| 1st gen. | 253.65 | 383.88 | 270.81 | 252.06 | 398.67 | 269.09 |
| 2nd gen. | 357.85 | 589.64 | 463.52 | 358.97 | 583.31 | 413.12 |

show that, unconditional on the educational choice, the averages of both abilities are higher for immigrants than for nonmigrants. Therefore, immigrants are positively selected with respect to both the abilities.

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.[37] 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 United States 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 they are 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 U.S. 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 $\xi$ 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 U.S. dollars per person and assume that the present value of lifetime earnings for high school educated nonmigrant Mexicans is 50,000 U.S. dollars. The first row shows the human capital of nonmigrant Mexicans, the second the first generation of immigrants, and the third row the second generation.

Compared to the benchmark case, in the counterfactual experiment the aggregate human capital per person of the second generation drops by 50,000 U.S. dollars from 463,000 to 413,000 U.S. dollars, more than 10%—Row 3, Columns 3 and 6. By sector, the difference is less pronounced and only appears for the college educated, who in the counterfactual scenario lose about 5,000 U.S. 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 implies 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.

[37] See the working paper version of this article, Caponi (2009, Appendix D) for a proof that the positive selection is due to altruism and is driven by the intellectual ability.

TABLE 9
COUNTERFACTUAL SIMULATION: POLICY EVALUATION—HUMAN CAPITAL

| | Benchmark | | | Counterfactual 1 | | | Counterfactual 2 | | |
|---|---|---|---|---|---|---|---|---|---|
| Generation | (H) | (C) | (A) | (H) | (C) | (A) | (H) | (C) | (A) |
| Nonmigrant | 50.00 | 172.85 | 68.99 | 49.97 | 173.30 | 69.69 | 49.97 | 170.86 | 65.18 |
| 1st gen. | 253.65 | 383.88 | 270.81 | 267.37 | 398.67 | 278.43 | 256.24 | 411.75 | 315.11 |
| 2nd gen. | 357.85 | 589.64 | 463.52 | 358.97 | 583.31 | 459.73 | 357.93 | 607.26 | 481.22 |
| Pol. cost | | | | 13.76 | | 12.60 | | 35.07 | 13.27 |

NOTES: Counterfactual 1: $z_H = 0.2867$, $z_C = 0.6444$; Counterfactual 2: $z_H = 0.3457$, $z_C = 0.4584$.

## 7. POLICY EVALUATION

In this section, I first look a 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 by lowering the cost of a college education. 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, 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 of 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.[38] 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 U.S. dollars per person and assume that the present value of lifetime earnings for high school educated nonmigrant Mexicans is 50,000 U.S. dollars. Rows 1, 2 and 3 show the human capital of nonmigrant 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 generations. Compared to the benchmark case, the aggregate human capital of the first generation increases from about 271,810 to 278,430 U.S. dollars, an increase of 7,620 U.S. dollars per person. This compares with a cost of the policy of about 12,600 U.S. 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 for 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 270,810 to 315,110 U.S. dollars per person, an increase of about 44,300 U.S. dollars. Compared to the 13,270 U.S. dollars of its cost, the improvement is much larger and implies a net gain.[39] Moreover,

[38] I first assume that an equivalent of about 13,760 U.S. 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.

[39] The 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 10
COUNTERFACTUAL SIMULATION: REDUCING THE COST OF COLLEGE—HUMAN CAPITAL

| | Benchmark | | | Lower College Cost MX | | | Lower College Cost US | | |
|---|---|---|---|---|---|---|---|---|---|
| Generation | (H) | (C) | (A) | (H) | (C) | (A) | (H) | (C) | (A) |
| Nonmigrant | 50.00 | 172.85 | 68.99 | 50.16 | 167.70 | 71.91 | 49.98 | 172.55 | 68.57 |
| 1st gen. | 253.65 | 383.88 | 270.81 | 253.35 | 387.21 | 267.19 | 254.14 | 379.09 | 273.49 |
| 2nd gen. | 357.85 | 589.64 | 463.52 | 358.30 | 580.92 | 457.65 | 360.03 | 581.34 | 465.04 |

the gain is transferred to the second generation as well, which earns 17,700 U.S. 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 nonmigrant 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 benefits not only 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 United States. 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 U.S. cost of college by a transfer of the same amount to every student attending a college in the United States.

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 United States, which goes from 270,810 of the benchmark to 267,190 under the first policy, a decrease of 3,620 U.S. dollars. The reduction is transferred to the second generation, which has a lower human capital per person, earning 5,870 U.S. dollars less than in the benchmark. In contrast, Column 1 show that the human capital possessed by Mexicans increases by 2,920 U.S. dollars from 68,990 to 71,910. Indeed, this would represent a significant gain for Mexico since the cost of the policy is only 645 U.S. dollars.[40] Overall, by effect of the policy there is a lower flow of migrants, especially among the most educated and with higher intellectual ability. Looking at Columns 2 and 4, in the 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. Overall, the policy is quite effective in containing the brain drain from Mexico to the United States

The second scenario shows the effects of a decrease in the cost of a college education in the United States. In this case the results are opposed to the ones reported above. The human capital

[40] Under the policy the percentage of college educated is 21.5% (see the working paper version of this article, Caponi, 2009, Table B.6 in Appendix B). Therefore, $0.215 * 3{,}000 = 645$.

per person in Mexico decreases, whereas it increases in the United States for the first and second generations. This is primarily due to the increase in the college share for the first and second generations and a reduction for the nonmigrant Mexicans. Overall, the increased skill premium in the United States 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 article provides an explanation of the intergenerational dynamics of earnings and educational attainment of successive generations of Mexicans in the United States. The explanation is based on three main concepts that are incorporated into my model. 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 upon migration. The second concept is that individuals are endowed with two abilities. The intellectual ability is used if some college education is acquired, and a manual ability is used otherwise. The third concept is that there is a transfer of abilities from parents to their children. In this respect, the estimation results suggest that immigrants' capacity to transfer their abilities to their children is not reduced. Therefore, although 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.

By allowing agents to be endowed with two distinct abilities, the model captures the complexity of the selection mechanism. The estimation results indicates that (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, I also find that even among the high school educated, immigrants are positively selected with respect to intellectual ability. Because immigrants care about their children, those with larger amounts of intellectual ability tend to migrate more, expecting their children to take full advantage of the inherited human capital. At the same time they also tend to remain high school educated to avoid the large loss of human capital associated with the use of the intellectual ability. 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 that is transferred to future generations.

In this sense, the findings in this article reverse the pessimistic view implied by the 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 a higher difficulty in adapting 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 no other obstacles that prevent this integration.

Finally, the article evaluates alternative policies aimed at integrating immigrants into the U.S. labor market or reducing the brain drain from Mexico to the United States. The simulation results show that policies that reduce the loss of human capital faced by college educated rather than by high school educated immigrants generate direct positive effects on the overall human capital and earnings of the first generation, but also indirect effects on the first and the second generations by strengthening the self-selection process of immigrants. In contrast, policies aimed at contrasting the brain drain from Mexico by reducing the cost of education is effective in improving the human capital distribution by increasing the educational attainment and lowering the incentive to migrate of the most talented Mexicans. It also reduces the amount of human capital brought by immigrants to the United States and reduces the positive selection.

## APPENDIX: DATA, MOMENTS, AND COVARIANCE MATRIX

A.1. *Data.* 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 United States of first, second, and third generations. Information about Mexican-born individuals living in the United States is obtained using the pooled 1994–2008 CPS data. To obtain an estimate of the Mexican population living in the United States in 2000, I reweigh 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 reweighting guarantees that the sum of all weighted observations from the pooled CPS data reproduce the Mexican population present in the United States 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 United States. 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 the case of first-generation immigrants, who entered the United States 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.[41,42]

The moments are based on the following regression model:

$$\begin{aligned}\log(w_i) = {} & \beta_0 + \beta_1 D_{1Hi} + \beta_2 D_{2Hi} + \beta_3 D_{3Hi} + \beta_4 D_{0Ci} + \beta_5 D_{1Ci} \\ & + \beta_6 D_{2Ci} + \beta_7 D_{3Ci} + \gamma X_i + \zeta_i,\end{aligned} \tag{A.1}$$

where $\log(w_i)$ represents log-hourly wages in 2000 U.S. dollars regressed on a set of dummy variables $D's$ 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 $D$ dummies to allow different effects for each group, dummies for survey years and for geography and, for data on Mexicans in the United States, the year of birth of the individual.[43] Dummy variables were constructed to represent the

[41] 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 in 2003. From 1996 the CPS does not set all the top-coded observations equal to the top-code level. Instead, the average 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 reassign 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.

[42] For the United States, 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.

[43] The geographical dummies, included to take into account the different purchasing power of earnings in different locations, are for states in Mexico and metropolitan status in the United States. The reference groups are large cities for the United States and the district of Mexico City for Mexico.

TABLE A.1
EARNINGS GAPS BETWEEN DIFFERENT GENERATIONS OF MEXICANS

| | Standard Definition | | Stricter Definition | |
|---|---|---|---|---|
| Dependent Var.: Log-Hourly Wage | Men | Women | Men | Women |
| Intercept | 0.6474 | 0.6039 | 0.6444 | 0.6105 |
| | (0.0129) | (0.0197) | (0.0124) | (0.0191) |
| 1st gen. imm. w/o coll. | 1.6263 | 1.3507 | 1.6231 | 1.3629 |
| | (0.0194) | (0.0274) | (0.0167) | (0.0249) |
| 2nd gen. Mexicans w/o coll. | 1.9742 | 1.6701 | 1.9676 | 1.6769 |
| | (0.0261) | (0.0332) | (0.0246) | (0.0322) |
| 3rd gen. Mexicans w/o coll. | 1.9642 | 1.6384 | 1.9636 | 1.6504 |
| | (0.0209) | (0.0272) | (0.0188) | (0.0254) |
| Mexicans with coll. | 1.2230 | 1.1440 | 1.2232 | 1.1440 |
| | (0.0081) | (0.0118) | (0.0079) | (0.0116) |
| 1st gen. imm. coll. | 2.0095 | 1.7161 | 2.0490 | 1.7552 |
| | (0.0276) | (0.0391) | (0.0236) | (0.0342) |
| 2nd gen. Mexicans coll. | 2.4288 | 2.1212 | 2.4065 | 2.1275 |
| | (0.0274) | (0.0332) | (0.0268) | (0.0331) |
| 3rd gen. Mexicans coll. | 2.3636 | 2.0889 | 2.3628 | 2.1015 |
| | (0.0218) | (0.0276) | (0.0197) | (0.0258) |
| $N$ obs. | 133,567 | 64,784 | 147,044 | 69,988 |
| $R^2$ | 0.6424 | 0.6192 | 0.6689 | 0.6315 |

NOTE: Standard errors in parentheses.

generations of Mexicans: $D_{iH}$ and $D_{iC}$ with $i = 0, 1, 2, 3$ for Mexicans living in Mexico, first, second, and third generation living in the United States with high school and college respectively. The omitted dummy captured by the intercept is $D_{iH}$.

Table A.1 reports the earnings moments resulting from the estimation of Equation (A.1).

A.2. *Moments.* First, I build a new earnings variable once I net out of the effect of other exogenous variables:

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

I also define $D_i = D_{iH} + D_{iC}$ with $i = 0, 1, 2, 3$ for all generations and $D_{[01]} = D_0 + D_1$ 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:

$$D_{1i} = R_m D_{[01]} + \epsilon_{mi}. \tag{A.3}$$

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

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

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:

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

TABLE A.2
ORIGINAL COVARIANCE MATRIX

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

*Assumptions about the error terms*. The basic assumption about the error terms is that they are i.i.d. 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*. See Table A.2.[44]

REFERENCES

ALTONJI, J. G., AND U. DORASZELSKI, "The Role of Permanent Income and Demographics in Black-White Differences in Wealth," *Journal of Human Resources* XL (2005), 1–30.

ATTANASIO, O., AND C. BINELLI, "Mexico in the 1990s: The Main Cross-Sectional Facts," *Review of Economic Dynamics* 13 (2010), 238–64.

BEHRMAN, J. R., AND M. R. ROSENZWEIG, "Does Increasing Women's Schooling Raise the Schooling of the Next Generation?" *American Economic Review* 92 (2002), 323–34.

BINELLI, C., "The Demand-Supply-Demand Twist: How the Wage Structure Got More Convex," Mimeo, Oxford University, 2009.

BORJAS, G. J., "Self Selection and the Earnings of Immigrants," *American Economic Review* 77 (1987), 531–53.

———, "The Intergenerational Mobility of Immigrants," *Journal of Labor Economics* 11 (1993), 113–135.

———, "The Economic Progress of Immigrants," in G. J. Borjas, ed., *Issues in the Economics of Immigration* (Chicago: University of Chicago Press, 2000), 15–49.

BRIGHAM, C. C., *A Study of American Intelligence* (Princeton, NJ: Princeton University Press, 1923).

CAPONI, V., "Heterogeneous Human Capital and Migration: Who Migrates from Mexico to the US?" IZA—Institute for the Study of Labor, Bonn, Germany (November 2006).

———, "Intergenerational Transmission of Abilities and Self Selection of Mexican Immigrants," Working Papers #2, Ryerson University, Department of Economics, 2009.

CARD, D., "Is the New Immigration Really so Bad?" *Economic Journal* 115 (November 2005), F300–23.

CHIQUIAR, D., AND G. H. HANSON, "International Migration, Self Selection, and the Distribution of Wages: Evidence from Mexico and United States," *Journal of Political Economy* 113 (2005), 239–81.

CHISWICK, B. R., "The Effects of Americanization on the Earnings of Foreign-Born Men," *Journal of Political Economy* (1978), 897–921.

CUNHA, F., AND J. HECKMAN, "The Technology of Skill Formation," *American Economic Review* 97 (May 2007), 31–47.

DUNCAN, B., AND S. J. TREJO, "Ethnic Identification, Intermarriage, and Unmeasured Progress by Mexican Americans," NBER Working Paper No. 11423, 2005.

GLICK, J. E., AND M. J. WHITE, "Post-secondary School Participation of Immigrant and Native Youth: The Role of Familial Resources and Educational Expectations," *Social Science Research* 33 (2004), 272–299.

GOURINCHAS, P.-O., AND J. A. PARKER, "Consumption over the Life Cycle," *Econometrica* 70 (2002), 47–89.

HECKMAN, J. J., AND J. SCHEINKMAN, "The Importance of Bundling in a Gorman-Lancaster Model of Earnings," *The Review of Economic Studies* 54 (1987), 243–55.

——, AND G. L. SEDLACEK, "Self-Selection and the Distribution of Hourly Wages," *Journal of Labor Economics* 8 (1990), S329–63.

——, J. STIXRUD, AND S. URZUA, "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," NBER Working Papers 12006, 2006.

KAO, G., AND M. TIENDA, "Optimism and Achievement: The Educational Performance of Immigrant Youth," *Social Science Quarterly* 76 (1995), 1–19.

KUROKAWA, Y., "Skill Intensity Reversal and the Rising Skill Premium: Evidence from the U.S. and Mexico," MPRA Paper 14013, University Library of Munich, Germany, July 2006.

LAIBSON, D., A. REPETTO, AND J. TOBACMAN, "Estimating Discount Functions with Consumption Choices over the Lifecycle," Mimeo, Harvard, 2005.

MAYER, A., "Education, Self-Selection, and Intergenerational Transmission of Abilities," *Journal of Human Capital* 2 (2008), 106–28.

MCFADDEN, D., "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration," *Econometrica* 57 (1989), 995–1026.

MCKEEVER, M., AND S. L. KLINEBERG, "Generational Differences in Attitudes and Socioeconomic Status among Hispanics in Houston," *Sociological Inquiry* 69 (1999), 33–55.

[44] Details about its derivation are available by the author upon request.

NEWEY, W. K., AND D. MCFADDEN, "Large Sample Estimation and Hypothesis Testing," in R. F. Engle and D. L. McFadden, *Handbook of Econometrics*, Volume 4 (Amsterdam: Elsevier, North-Holland, 1994), 2111–245.

SATTINGER, M., "Assignment Models of the Distribution of Earnings," *Journal of Economic Literature* 31 (1993), 831–80.

WILLIS, R. J., AND S. ROSEN, "Education and Self-Selection," *Journal of Political Economy* 87 (1979), S7–36.