Are Immigrants Positively or Negatively Selected?

The Role of Immigrant Selection Criteria and Self-Selection

by

Abdurrahman B. Aydemir*

Family and Labour Studies Division

Statistics Canada

March 2003

Abstract

This paper specifies and estimates a structural model of international migration using micro data. This provides a direct test of human capital theory that suggests that individuals respond to the earnings differentials across countries while making their migration decisions. The paper specifies migration as a joint outcome of two decision makers, i.e. the individual who decides to apply for migration and the host country that reviews applications, and identifies the factors determining the decision of these two players. The empirical results provide evidence in support of the human capital model. It is also shown that both the host country and the individual have significant impacts on the resulting characteristics of immigrants. The results suggest negative self-selection at the application stage both in terms of observed and unobserved characteristics and a positive selection at the review step by the host country. Although there is negative self-selection in terms of schooling among applicants, as a result of the positive selection at the review step the resulting migrants are positively selected. However, in terms of unobservable characteristics the review step is unable to reverse the negative self-selection that occurs at the application stage, and the resulting migrants are negatively selected in this dimension.

Key Words: immigration, self-selection, Roy Model, immigration policy
JEL: J61, J68

* I would like to thank Barry Chiswick and Chris Robinson for their helpful comments. I am solely responsible for the remaining errors in the paper. The views in this paper belong to the author and does not necessarily represent the views of Statistics Canada. Correspondence: aydeabd@statcan.ca

1. Introduction

The literature studying international migration recognizes that immigrants arriving at various host countries are not randomly selected from the populations of the sending countries. The human capital theory postulates that these individuals are self-selected as they respond to earnings differentials between the source country and the host country. There is a large literature based on this working assumption. Although this key assumption has been tested directly in internal migration literature by constructing the earnings differences for movers and stayers, such a test has not been done in the international migration context. Also in the international migration context while it is recognized that immigrants are a self-selected sample, both the theoretical and the empirical literature ignores the selection of immigrants by the immigrant receiving countries. What determines the resulting selectivity of immigrants, both in terms of observed and unobserved characteristics, is this double selection process. This paper tests the human capital theory in the international context and provides evidence about the nature of selection in terms of observed and unobserved characteristics of the immigrants. The paper distinguishes between the self-selection of immigrants which forms the applicant pool for migration and the selection by the host country which determines who is accepted among the applicants. The separate effects of this two-step selection process on the characteristics of immigrants are identified.

The difficulty with the testing of the human capital model in this context lies primarily in the need to construct counter-factual earnings estimates, i.e. what the earnings of an immigrant would have been had that individual not migrated, and vice-versa for non-migrants. Secondly, in any given data set, contrary to the internal migration process, it is possible to identify either only immigrants, as in the case of a Census from a host country, or only non-migrants as in the case of a Census from a sending country. Hence, to study the international migration process it becomes necessary to use data from both the source and the host countries. Thirdly, immigrant receiving

countries has very important effects on the resulting selectivity of immigrants through their immigration policies. This feature of the process needs to be incorporated in the empirical testing. For example, Canada and Australia selects their skilled immigrants through a points test. Among the applicants only those who posses the characteristics that earn them sufficient number of points are admitted while the rest are rejected. In the United States special visas are allocated to different industries that have demand for skilled workers and employers bring in skilled immigrants through these visas. European countries such as Germany used 'guest-worker' programs to bring in workers to be employed in certain industries most of them becoming permanent immigrants. Therefore what determines the resulting selectivity of the international migration process is not only the supply-side which determines the self-selection as recognized by the literature but also the visa rationing by the host country.

Among the studies of internal migration Robinson and Tomes (1982) is the first paper to test the responsiveness of individual migration to the earnings differentials by constructing opportunity wages of individuals; that is the wages of migrants had they stayed and of stayers had they moved. Their empirical results offer strong support for the human capital model. There are only few studies studying the determinants of international migration (Lucas, 1985, Taylor, 1986, O Grada, 1986, Stark and Taylor, 1991, and Adams, 1993). These studies use data sets collected from the immigrant sending countries, generally covering small areas, which can identify individuals who emigrated from the country. These studies look at the relationship between the migration decision and several individual and household characteristics such as age, schooling, marital status, home ownership, household income and household size. While providing important insights about the correlation between such characteristics and the outcome of migration, these studies are limited in a number of ways. Firstly, immigration policies of the host countries are completely ignored. This prevents estimation of the structural parameters of

interest underlying individuals' desire to migrate and also any analysis of the immigration policy. Secondly, the data used in these studies are generally collected from small areas that generated emigrants in the past and are not representative of the source country. While some of this data have information about the household income in the source country, they don't have any information about post migration outcomes such as the host country earnings or income of the individuals that migrated. To my knowledge there is no study in the international context that uses nationally representative data sets for both the source and host countries and tests the human capital model.

This paper tests the effects of earnings differentials between the host country and source country by constructing the opportunity wages similar to Robinson and Tomes (1982). Moreover, in doing so the paper explicitly models the selection of individuals by the host country and hence the effect of immigration policy. Estimation of the model and the opportunity wages also provide important insights about the selectivity of observed and unobserved characteristics of immigrants. The selectivity of immigrant characteristics has been the subject of a series of papers in the literature. Early literature on international migration assumes a favorable selection among immigrants in terms of both observed and unobserved characteristics. This positive selection argument is used to explain the finding in cross-sectional studies of immigrant earnings that age-earnings profile of immigrants crosses the age-earnings profiles of natives with same observed characteristics after ten to fifteen years after migration. This is explained by the unobserved characteristics of immigrants who may be more able and more highly motivated (Chiswick, 1978). This view has been challenged by Borjas (1987) who developed a model of immigrant self-selection based on the model of Roy (1951) where the focus is on selectivity in unobserved characteristics[1]. The theoretical model demonstrates that immigrants may not

[1] Borjas (1991) extends this framework and discusses selection in terms of observed characteristics.

necessarily be positively selected as is commonly assumed. The Roy model suggests that country-specific characteristics of the income distribution (and mobility costs) determine the unobserved quality of immigrants. Macro level data reflecting the economic and social conditions of the sending countries relative to the host country (U.S.) at the time of migration is included in the estimated equations of immigrant-native entry wage differentials. The key variable is the one measuring the income inequality and is calculated as "The ratio of household income of the top 10 percent of households to the income of the bottom 20 percent of the households" (Borjas, 1987). Borjas argues that if earnings between the host country and the source country are positively and strongly correlated, positive selection in terms of unobserved characteristics is observed whenever the host country has more income inequality then the source country and negative selection is observed otherwise. The empirical testing of this model, however, resulted in mixed results and led to some controversy in the literature.[2] Although the empirical results provide some evidence supporting the model, the variable measuring the relative income inequality doesn't always have the expected sign or the statistical significance (Borjas, 1987, Jasso and Rosenzweig, 1990, Borjas, 1990, Borjas, 1991, Cobb-Clark, 1993). In the empirical studies of immigrant selectivity the list of and motivation for the macro level variables are necessarily ad hoc which leads to much of the criticism and the mixed results (Cobb-Clark, 1993). Chiswick (1999, 2000) argues that the measure used for household income inequality may be poorly related to the relevant variable of relative skill differentials. The primitive parameters in the Roy model that determine selectivity in terms of unobserved skills describe the dispersion in opportunities for a given level of socioeconomic characteristics. Borjas (1987) admits that the income inequality in the income distributions used in the empirical analysis do not correspond exactly to those parameters as they are affected by a host of other

[2] See Borjas (1987), Jasso and Rosenzweig (1990), Borjas (1990), Borjas (1991), Cobb-Clark (1993) and Chiswick

factors. Jasso and Rosenzweig (1990) further argue that results of Borjas (1987) are sensitive to the choice of source countries included in the analysis. Greenwood and McDowell (1991), and Cobb-Clark (1993) note that immigration policy is not taken into account in this literature and they include variables in their reduced form equations on an ad hoc basis trying to capture these effects. Their results suggest that immigration policy is important in determining observed outcomes.[3]

Estimation of the model of this paper along with the opportunity wages provides direct evidence on selectivity of observed and unobserved characteristics. Moreover, the paper goes one step further and highlights the importance of selection by the host country which has been largely ignored in the literature. The empirical results in this paper support the human capital model with the evidence showing that wage differentials between the host and the source country affect the migration decision in the predicted direction. This is the first direct evidence in international migration context in favor of the human capital model. The results also show that both the host country and the individual each have significant impact on the outcome of international migration. Finally, estimation of the model, wage equations and predicted earnings of immigrants in the source country provide evidence on selection issues. The results suggest negative self-selection at the application stage both in terms of observed and unobserved characteristics which is consistent with the predictions of the Roy model. However, when the predicted earnings of immigrants in the source country are compared to the predicted earnings of non-migrants it is shown that they come from the upper end of the income distribution. This result is due to positive selection that occurs at the review step by the host country which shows

---

(1999, 2000)

[3] Borjas (1987) assumes that visa restrictions enter into the equation by increasing the costs of migration. The variable used to capture the costs of migration is the distance between the host country and the source country which assumes these costs are fixed and same for all individuals from a given source country. Note, however, that these "additional costs" due to selection by the host country will be different among individuals of a given source country

the importance of rationing of visas by the source country in determining the resulting selectivity of immigrants.

The next section outlines the theoretical model and estimation issues. Section 3 describes the data used in the analysis and section 4 presents the results of the estimation. Section 5 concludes.

2. Model

This paper estimates a structural model of international migration based on the framework developed by Aydemir (2003). In this model migration is specified as a two step process:

Step (1): **Application**: Potential migrants apply to migrate to a host country.

Step (2): **Review**: The host country chooses migrants from the applicant pool.

The two decision makers, the individual and the host country, determine the outcome of international migration. The individual makes a decision on whether to apply or not, and the host country decides whether to accept a given application or reject it.

The individual *i* located in the source country (*country 2*) and considering migration to the host country (*country 1*) will apply if the expected gain from application exceeds the application cost, C_{ai} :

$$P_i^a\ (\ w_{1i} - C_{mi}) + (1 - P_i^a)\, w_{2i} - w_{2i} > C_{ai} \qquad (1)$$

where w_{1i} and w_{2i} are the present value of life time earnings in the host and the source countries, P_i^a is the probability that application will be accepted and C_{mi} is the moving costs. The expected income if the individual applies is given by $P_i^a\ (\ w_{1i} - C_{mi}) + (1 - P_i^a)\, w_{2i}$ whereas the expected

and will lead to the selection among individuals. As Borjas focuses on unobserved quality differences across

income if the individual does not apply is simply, w_{2i} and the difference between the two give the expected gain of application.

If $C_{ai} = 0$, the criterion reduces to:

$$(w_{1i} - w_{2i} - C_{mi}) > 0 \qquad (2)$$

for non-zero P_i^a. The cost of application C_{ai}, is incurred at the time of the application. This is a variable cost depending on the number of dependents and does not include costs of moving, foregone earnings etc. which are included in C_{mi}. C_{ai} is generally small relative to the lifetime w's and is assumed to equal zero in the rest of the analysis.[4]

Following the previous literature, let[5,6]

$$C_{mi} = c_{mi}\, w_{2i}$$

where $c_{mi} \geq 0$ and using the approximation $\ln(1+c) \approx c$ (for small values of c), the equivalent criterion for applying becomes:

$$\ln w_{1i} - \ln w_{2i} - c_{mi} > 0 \qquad (3)$$

Let the present value of log lifetime earnings in the two countries be represented by the following linear functions:

---

countries, he ignores this important selection issue among individuals within a given source country.

[4] The assumption of $C_{ai} = 0$ rules out any effect of review by host country on application decision. Relaxation of this independence assumption can only be done with more information on costs and changes in the model, and is left for future research.

[5] Robinson and Tomes (1980) give theoretical and empirical justification for this assumption. One justification is that the cost of migration is proportional to the source country permanent income due to the existence of source country specific investment and/or endowments. The linear specification $C_{mi} = c_{mi}\, w_{2i}$ follows from the convenience of functional form. Note that the factor of proportionality c_{mi} is indexed by *i* allowing the costs to vary across individuals.

[6] Note that with the assumption that migration costs are proportional w_{2i} it is implicitly assumed that there are no direct (or out of pocket) costs. These costs, such as the cost of an airfare, are likely to be very small in the Canada-US context being studied here. Under no direct cost assumption, Chiswick (1999, 2000) shows the Roy model to be a special case of the human capital model.

$$\ln w_{1i} = X_{1i}\beta_1 + \varepsilon_{1i} \qquad (4.1)$$

$$\ln w_{2i} = X_{2i}\beta_2 + \varepsilon_{2i} \qquad (4.2)$$

And let

$$c_{mi} = X_{3i}\beta_3 + \varepsilon_{3i} \qquad (5)$$

where the X's are observable and the ε's are unobservable to the econometrician. The β coefficients represent the structural parameters of interest in the *application step*.

Next, define the following index:

$$I^*_{1i} = X_{1i}\beta_1 + \varepsilon_{1i} - X_{2i}\beta_2 - \varepsilon_{2i} - X_{3i}\beta_3 - \varepsilon_{3i} \equiv W_i\pi + \varepsilon_i \qquad (6)$$

so that the criterion for step (1) is $I^*_{1i} > 0$, or $\varepsilon_i > -W_i\pi$. The coefficients π represent the reduced form parameters of interest in the application step.

The other decision maker, i.e. the host country, accepts an application if

$$I^*_{2i} = X_{4i}\beta_4 + \varepsilon_{4i} > 0 \qquad (7)$$

where X_{4i} are observable characteristics of the potential immigrant and any measurable policy parameters and ε_{4i} are unobservables to the econometrician. In Canada immigrants other than those accepted under family class or due to humanitarian reasons (refugee class) are subject to a points test. In the point system several characteristics are assessed and given points according to a schedule, and admission depends on having sufficient number of points over the pass mark.[7] A subjective assessment by the immigration officer and unobserved factors such as a health problem that may cause rejection contributes to the error component.

[7] Appendix A provides further information about the Canadian Immigration Policy and the point system as of 1986-1990 period.

2.1 Estimation of the Parameters of the Two-Step Model

Estimation of the parameters from the reduced form index π and β_4 is called reduced form estimation of the 2-step model. Estimation of the parameters of the application step β_1, β_2 and β_3, along with β_4 is referred to as "structural estimation". The aim of the structural estimation is to obtain consistent estimates of β_1, β_2 that will be used to get consistent estimates of $\ln w_{1i}$ and $\ln w_{2i}$. Given the consistent predictors for $\ln w_{1i}$ and $\ln w_{2i}$, these may be used in the structural index:

$$I^*_{1i} = (X^{\wedge}_{1i}\beta_1) - (X^{\wedge}_{2i}\beta_2) - X_{3i}\beta_3 - \varepsilon_i$$

The probability of applying for migration for an individual is then given by:

$$\Pr\left\{\left[(X^{\wedge}_{1i}\beta_1) - (X^{\wedge}_{2i}\beta_2) - X_{3i}\beta_3\right]/\sigma_\varepsilon > \frac{\varepsilon_i}{\sigma_\varepsilon}\right\} \qquad (8)$$

On the other hand the probability that an individual i will be accepted (i.e. $I^*_{2i}>0$) is given by:

$$\Pr\left\{X_{4i}\beta_4 + \varepsilon_{4i} > 0\right\} \qquad (9)$$

Since we only observe whether an individual is a migrant or not, but don't observe the application step or application review step by the host country, the parameters of the reduced form (π and β_4) or the structural parameters (β_1, β_2 and β_3, along with β_4) can not be estimated by two separate probit models for (8) and (9) above. For example, in the data those who applied but were turned down are not known. The only information available is who actually migrated (that is who applied and was admitted) and who did not which leads to the partial observability framework of Poirier (1980). The model parameters can be obtained by estimating the following bivariate probit model:

$$Z_i = I_{1i} * I_{2i} \qquad i = 1,2,...,n$$

$Z_i = 1$ if individual i is observed as a migrant; 0 otherwise. Thus Z_i is the observed migration variable and the probability of migration is the joint probability:

$$\Pr(I_{1i} = 1, I_{2i} = 1) = \Pr(I_{2i} = 1 | I_{1i} = 1)\Pr(I_{1i} = 1) = \Pr(Z_i = 1)$$

The probability distribution of Z is given by:

$$\Pr(Z_i = 1) = \Pr(I_{1i} = 1 \text{ and } I_{2i} = 1) \equiv P_i^m$$
$$\Pr(Z_i = 0) = \Pr(I_{1i} = 0 \text{ or } I_{2i} = 0) = 1 - \Pr(I_{1i} = 1 \text{ and } I_{2i} = 1) = 1 - P_i^m$$

Given a random sample of observations on Z, the log likelihood for the sample can be specified given distributional assumptions on the ε's. Assume that each individual in the source country draws a realization of the pair $(\varepsilon, \varepsilon_4)$ from the bivariate normal distribution $g(\varepsilon, \varepsilon_4, \rho)$, where ρ is the correlation between ε and ε_4. For reduced form parameters of interest, given a random sample of observations $(Z_i | W_i, X_{4i})(i = 1,2,...,n)$, the log likelihood is given by:

$$L(\pi, \beta_4, \rho) = \sum_{i=1}^{n} \left\{ Z_i a_i^m \ln G(W_i\pi, X_{4i}\beta_4; \rho) + (1 - Z_i) a_i^{nm} \ln\left[1 - G(W_i\pi, X_{4i}\beta_4; \rho)\right] \right\} \qquad (10)$$

Letting $\theta = [\pi', \beta_4', \rho]$, the parameter vector θ is to be estimated.

Under the usual identification conditions (e.g. for the reduced form parameters π and β_4, one of the vectors $(W_i \text{ or } X_{4i})$ excludes at least one exogenous variable appearing in the other vector) and subject to a normalization rule in each equation, the parameters are identified (Heckman, 1976, 1978; Amemiya, 1978).[8],[9]

[8] Choice-based sampling is relevant in this study for the subpopulations of those who migrate and those who do not. The model is estimated along the method of Manski and Lerman (1977) by using the appropriate weights for choice-based sampling. The weight for a migrant (a_i^m) is equal to the fraction of migrants in a certain group (such as males, aged 18 to 65) in source country population over the fraction of migrants in the sample (with previous restrictions) used for estimation. The weight for a non-migrant (a_i^{nm}) is calculated similarly. The data provide the number of migrants from a source country over a period of time. Given the size of the source country population the weights can be calculated.

[9] Choice-based sampling also requires a correction for the standard errors. The correct standard errors for the choice-based sampling are calculated and reported later in the text with the parameter estimates.

A 3-stage estimation procedure is used to obtain consistent estimates of structural parameters. This method is an extension of the Heckman 2-stage method for sample selectivity. The three stage estimation procedure includes first, the estimation of the reduced form 2-step model that yields estimates of (π and β_4).

Second stage involves getting the consistent parameter estimates of the earnings equations to construct the earnings differentials $\left\{(\hat{X_{1i}}\beta_1)-(\hat{X_{2i}}\beta_2)\right\}$. The earnings functions cannot be estimated for all individuals originally resident in the source country. For the migrants we observe $\ln w_{1i}$ and for the stayers we observe $\ln w_{2i}$. Estimation of the earnings equations requires taking into account the double selection process. Equation (4.1) is estimated on the sample of migrants:

$$\begin{aligned} E\left[\ln w_{1i}\left|Z_i=1\right.\right] &= E\left[\ln w_{1i}\left|I_{1i}^*>0 \quad and \quad I_{2i}^*>0\right.\right] \\ &= X_{1i}\beta_1+E\left[\varepsilon_{1i}\left|I_{1i}^*>0 \quad and \quad I_{2i}^*>0\right.\right] \\ &= X_{1i}\beta_1+\rho_{13}^1\lambda_{41}^i+\rho_{23}^1\lambda_{42}^i+v_{1i} \end{aligned} \qquad (11.1)$$

Equation (4.2), on the other hand, must be estimated on the sample of stayers:

$$E\left[\ln w_{2i}\left|Z_i=0\right.\right]=X_{2i}\beta_2+E\left[\varepsilon_{2i}\left|Z_i=0\right.\right]$$

where

$$\begin{aligned} Z_i=0 \quad if \quad & \left(I_{1i}^*<0 \quad and \quad I_{2i}^*>0\right) \quad or \\ & \left(I_{1i}^*>0 \quad and \quad I_{2i}^*<0\right) \quad or \\ & \left(I_{1i}^*<0 \quad and \quad I_{2i}^*<0\right) \end{aligned}$$

Estimation of earnings equation for source country using the sample of stayers is then given by

$$\begin{aligned} E\left[\ln w_{2i}\left|Z_i=0\right.\right] &= X_{2i}\beta_2+E\left[\varepsilon_{2i}\left|Z_i=0\right.\right] \\ &= X_{2i}\beta_2+\rho_{13}^0\underbrace{\left(\lambda_{11}^*+\lambda_{21}^*+\lambda_{31}^*\right)_i}_{\lambda_1^*}+\rho_{23}^0\underbrace{\left(\lambda_{12}^*+\lambda_{22}^*+\lambda_{32}^*\right)_i}_{\lambda_2^*}+v_{2i} \end{aligned} \qquad (11.2)$$

Using the parameter estimates from the reduced form two-step model the sample selection terms in the two earnings equations (11.1) and (11.2) are constructed. Estimation of (11.1) and (11.2) provides consistent parameter estimates of β_1 and β_2 in (11.1) and (11.2) which are then used to construct the predicted earnings differentials.

Finally in the last stage of estimation, using these predicted earnings differentials and other relevant variables we estimate the bivariate probit model that corresponds to the following log likelihood function[10]:

$$L(\pi, \beta_4, \rho) = \sum_{i=1}^{n} \left\{ \begin{array}{l} Z_i a_i^m \ln G\left(1/\sigma_\varepsilon \left[(\hat{X_{1i}}\beta_1) - (\hat{X_{2i}}\beta_2) \right], X_3\beta_3, X_{4i}\beta_4; \rho\right) + \\ (1 - Z_i) a_i^{nm} \ln\left[1 - G\left(1/\sigma_\varepsilon \left[(\hat{X_{1i}}\beta_1) - (\hat{X_{2i}}\beta_2) \right], X_3\beta_3, X_{4i}\beta_4; \rho\right)\right] \end{array} \right\} \qquad (11.3)$$

Note that the probability of observing an individual as a migrant is given by

$$\Pr\left(\left[(\hat{X_{1i}}\beta_1) - (\hat{X_{2i}}\beta_2) - X_{3i}\beta_3 \right] / \sigma_\varepsilon > \frac{\varepsilon_i}{\sigma_\varepsilon} \middle| X_{4i}\beta_4 > -\varepsilon_{4i} \right) * \Pr\left(X_{4i}\beta_4 > -\varepsilon_{4i} \right)$$

The coefficient on the predicted earnings differential $\left\{ (\hat{X_{1i}}\beta_1) - (\hat{X_{2i}}\beta_2) \right\}$ is given by $1/\sigma_\varepsilon > 0$. This restriction may be tested by estimating the bivariate probit (11.3).

The earnings equation (11.1) is estimated for migrants while (11.2) is estimated using a sample of non-migrants. Using the reduced form parameter estimates, sample selectivity terms λ_{4j}^i for migrants, and λ_j^* for non-migrants are calculated for inclusion in the wage generating functions. The coefficients obtained on λ_{4j}^i and λ_j^* provide information on whether there is positive or negative selection in migrant or non-migrant categories. For ($j = 1, 2$), coefficients on λ_{4j}^i indicate whether a migrant would earn in the host country, other things being equal, more than the average taken over both migrant and non-migrants. Similarly, the coefficients on λ_j^* for

$j = 1, 2$ is interpreted as whether the non-migrants would earn more in the source country, other things being equal, than the average taken over both migrants and non-migrants.

2.2 Empirical Specifications of the Equations

The earnings functions are specified as semi-logs proposed by Mincer (1974):

$$\ln w_{ji} = X_{ji}\beta_j + \varepsilon_{ji}$$

where

$$X_{ji} = \{\text{schooling, degree, experience, language, training}\}_i$$

and higher order terms in experience. The language variable is in general important in immigration context. In this study, however, where migration form the U.S. to Canada is concerned all individuals in the data report being either native English speakers or speak English very well. Therefore language variable is omitted. Also, lack of information on training in the data sets leads to omission of this variable.

The variables ε_{1i} and ε_{2i} include general unobserved ability and unobserved country specific capital. It is assumed that, over the entire population of individuals initially located in the source country, ε_{1i} and ε_{2i} have zero means, variances σ_{11}, σ_{22} and covariance σ_{21}. No restrictions are imposed on the sign of σ_{21}. The parameters β_1 and β_2 are not constrained to be equal, allowing rates of return for each characteristic to vary by country. For example, the location where the highest degree and training are completed can be important in the immigration context.

The factor of proportionality c_{mi} is given by: $c_{mi} = X_{3i}\beta_3 + \varepsilon_{3i}$ where

[10] There is a technical appendix for the details of this three stage procedure and the earnings equations (11.1) and (11.2) available from the author upon request. The framework for earnings equations follows and extends Tunali (1986).

$$X_{3i} = \{\text{family size, language, marital status, schooling}\}_i$$

$$\varepsilon_{3i} = \{\text{unobservable cost components}\}_i$$

ε_{3i} is assumed to have zero mean, variance σ_{33} and covariances σ_{23}, σ_{13}. It is argued that the presence of children, due to changing schools, etc., inhibits migration. Also, there is empirical evidence that the existence of a spouse increases costs of migration (Mincer, 1978). Schooling and language are hypothesized to reduce the information costs via better information and job prospects and lower the cost of moving. Again in the U.S.-Canada context the language variable is dropped.

Given the above specifications of X_{1i}, X_{2i}, X_{3i} and after replacing experience with age, in the reduced form estimation W_i in (6) is given by:

$$W_i = \{\text{schooling, degree, age, family size, marital status}\}_i$$

The vector X_{4i} in the second selection index is given by:

$$X_{4i} = \{\text{schooling, age, occupation, degree}\}_i$$

$$\varepsilon_{4i} = \{\text{unobservable components affecting admission}\}_i$$

These variables in X_{4i} are the characteristics evaluated in the point system. More educated individuals get higher points. In general, applicants in white-collar occupations (Executive and Administrative, Professional Specialty, Teaching and Related Occupations) are awarded higher points compared to those in sales and blue collar occupations. Younger applicants get higher points under the age factor, whereas those more experienced are awarded higher points under the experience factor.[11] Also, those who have arranged employment are more likely to be admitted as they are awarded extra points. It is assumed that the probability of having arranged

[11] An identification problem arises if experience, age and schooling are entered simultaneously since experience is calculated as (age - years of schooling – 6).

employment depends on occupation, experience and degree. Unobservable components affecting admission includes the points under the 'personal suitability' factor which is determined by the immigration officer at an interview. Also, other factors such as a medical condition preventing admission of an applicant or the factors that affect arranged employment that can not be captured by the characteristics in X_4, such as motivation, contribute to the error term. The mean of the personal suitability points is captured in the intercept term and the remaining error component is assumed to have zero mean and constant variance over the population of the source country.[12]

3 Data

Immigrant sample used in the estimation is obtained by matching the Landings Records (LIDS) with the Immigration Database (IMDB).

LIDS data contains demographic data (e.g. age, marital status, last permanent residence), program data (immigrant category, special program codes, principal applicant code) and personal characteristics (intended occupation, years of schooling, level of education, self-assessed knowledge of an official language). These data are recorded as of the date of issue of the landing visa and are available for all landings in Canada. Longitudinal IMDB data matches the landings records with the earnings information in tax files providing several observations on earnings of an immigrant.[13]

---

[12] Above specification assumes that the selection index I^*_{2i} is that ε_4 is not correlated with X_4. Potentially a non-zero correlation between ε_4 and X_4 may exist if the migration officers decision for points under "personal suitability" depends on the points obtained from other factors. Using the distribution of point scores, the correlation between ε_4 and individual components of X_4 as well as the total number of points obtained from X_4 are investigated and no significant correlation is found.

[13] To be included in the IMDB a migrant has to file at least one tax return starting from the date of application. A study done by Carpentier and Pinsonneault (1994) finds that the IMDB is representative of the tax-filing population. Average characteristics in the IMDB resemble those generally found in the labour force.

This study also uses a sample from the Immigration Data System Overseas (IDSO) which provides the points awarded to principal applicants for each characteristic assessed under the point system. IDSO also provides other administrative information such as the date the application was received, and the date of landing. The information in IDSO is used for testing the assumptions regarding the error term ε_4 as discussed in the previous section.

The distinguishing features of the data on migrants are the inclusion of immigrant class/category, the special program codes and the principal applicant codes. Information on immigrant class (e.g. refugee class, independent class etc.) determines the selection criteria that apply to an immigrant. Principal applicant information indicates which immigrants are assessed and which are dependents. Special program codes indicate cases admitted under relaxed criteria. These variables, which does not exist in other data sources used for studying immigration such as the Census, are crucial in determining which individuals goes through the points test. This additional information allows for a detailed and more careful analysis of the immigration policy and labour force outcomes of immigrants.

The sample of immigrants consists of male immigrants aged 18-65 who applied and were accepted from the U.S. between 1986 and 1990. There were several changes in the point system between 1980 and 1990. Over the period of 1986 to 1990 the selection criteria did not change for the independent class allowing the study of migrants admitted under a constant policy regime. The migrant samples are further restricted to skilled workers admitted under the independent class, who are principal applicants. These are individuals who are seeking admission on the basis of their labour market skills by going through the points test.[14] Immigrants whose intended

[14] Immigrants can enter Canada under Family or Assisted Relative classes or the Refugee class. While illegal immigration is an important issue for other countries e.g. the U.S., illegal immigration and refugee migration is not an important issue between Canada and the U.S. Although the decision of individuals under which immigrant class to apply is interesting, it is beyond the scope of this paper. Also note that, as Chiswick (1999, 2000) argues the

destination is Quebec are excluded since the province of Quebec has its own selection criteria different from the rest of Canada. Immigrants who are admitted under special programs or those that pass the point test only with the immigration officer's discretion are also excluded[15]. The resulting migrant sample from the U.S. consists of 2,500 records.

The non-migrant sample is a 1/60 random sample from the 1/100 U.S. Census restricted to males aged 18 to 65, in civilian labour force and not in school at the time of the census. This sample refers to non-migrants, that is individuals who didn't apply or who applied but were rejected for immigration to Canada. The resulting non-migrant sample from the U.S. consists of 10305 records. Pooling sample of migrant and non-migrants consist of 12,805 records.

$\ln WAGE$ obtained from the 1990 US Census includes wages and salaries and income from self-employment and the figures are first converted to 1990 constant US dollars and then to 1990 Canadian dollars using purchasing power parity (PPP) index[16]. Given the cross-sectional nature of the data, it contains only one observation on the age-earnings or experience-earnings profile of each individual. Given this restriction, the model characterizes an individual's profile in the source country by using a single variable (permanent income) as described above. One could also compute the lifetime earnings of an individual using the lifetime-wage profile implied by the cross-sectional data. However, given the likely presence of cohort effects this can lead to a biased estimate of life-time earnings.

---

favorable selection is expected to be strongest among individuals migrating under skilled class relative to other immigrants e.g. refugees who are primarily moving for non-economic reasons.

[15] Independent class applicants may be accepted even if they don't have enough points, by way of the immigration officer's discretionary power. This discretion, which is exercised very rarely, is different from the immigration officer's regular evaluation for personal suitability. An example of a special program would be accepting immigrants from a country that had a natural disaster. For the admission of these migrants some or all of the characteristics in X_4 may be irrelevant, and these migrants are excluded from the sample as well.

[16] CPI figures are obtained from: Consumer Price Indexes, Sixteen Countries, 1950-1998, US Department of Labor, Bureau of Labor Statistics, Office of Technology, June 8, 1999

PPP figures are obtained form: Cansim Purchasing Power Parity, System of National Accounts Classification, Expenditure Based, Statistics Canada

The same definition is used for the $\ln WAGE$ for earnings observations of immigrants from the IMDB as in the US data. IMDB provides multiple observations on earnings for a given individual. For example, an immigrant who arrived in 1986 may have up to nine observations on earnings for years 1987 to 1995[17]. All earnings figures are expressed in 1990 constant Canadian dollars using the CPI index. Using each tax year as a cross-section and pooling the earnings data over 1987 to 1995, the wage function (11.1) is estimated. Having more than one observation for a given individual provides better estimates of host country earnings equation as recent immigrants are expected to earn less than earlier immigrants with the same characteristics.

Tables 1 give the definitions of variables while Tables 2 presents the descriptive statistics for migrants and non-migrants. Table 2 reveals that migrants are composed of younger, more educated individuals concentrated in white-collar occupations that are more likely to be married relative to non-migrants.

4 Estimation Results

4.1 Results from the Reduced Form Estimation

Table 3 presents the results from the estimation of the reduced form 2-step model of migration. The coefficients in the first selection index refers to reduced form coefficients π in equation (6) indicating the total effect of the exogenous variables on the probability of applying, acting through the wage differential for migrating versus staying and through the cost of moving. The coefficients of the second selection index refer to β_4 in (7) and give the effect of exogenous variables on the probability of acceptance.

The important result that emerges from table 3 is that individuals with higher education are less likely to apply while they are more likely to be accepted through the point system. This

[17] Earnings data for those whose year of arrival is the same as the tax year is not used in the estimation since they

result shows that individuals in the U.S. are negatively self-selected at the application stage in terms of observable characteristics. However, there is a positive selection at the review step by the host country among those who apply. The importance of taking into account the selection by the host country is illustrated by Aydemir (2003) which estimates the reduced form 2-step model presented here as well as a reduced form model that ignores the selection by the host country. When estimation is done for two source countries, the U.S. and U.K., the results from the reduced form model that ignores the selection by the host country shows that more educated individuals are more likely to be observed as immigrants. The pervious literature ignoring the selection by the host country would interpret these results as evidence of positive self-selection where more educated have more incentives to migrate and hence more willing to become immigrants. When reduced form 2-step model is estimated, however, the results show that while more educated individuals from UK have more incentives to migrate, the reverse is true for individuals in the US.

Other results in table 3 referring to first selection index show that having more dependents decreases the probability of application for migration through increasing costs of migration, increasing age initially increases the probability of applying for migration, later decreases this probability. Results from the second selection index are in general consistent with the structure of the point system where older individuals are less likely to be accepted, whereas those in white collar occupations who get higher points under occupation factors are more likely to be admitted. Hence, the results for education variable suggest negative self-selection at the application stage and positive selection at the review step by the host country. These results from the reduced form 2-step model are used to construct the sample selectivity terms to get consistent parameter estimates in (11.1) and (11.2) and results are presented in the next section.

---

may not have spent a full year in the country after their arrival.

4.2 Results from the Estimation of the Wage Equations

The results of the estimation of the earnings equations for the host country and the source country are presented in tables 4 and 5. Schooling and experience have the expected signs in both equations with the source country (US) returns to these characteristics being larger than the returns in the host country (Canada). Since the earnings equation results will be used to predict the counterfactual earnings (e.g. results from the estimation of the host country earnings equation will be used to predict earnings of non-migrants in Canada had they migrated) certain variables such as region of residence are not included in the regression. If such a variable was included we would have to know the region of residence in Canada a non-migrant would be in had that individual migrated. Instead, the simple specification allows us to average over regions in this example, thus enabling us to predict the counter factual earnings of a non-migrant.

Of particular interest here are the estimates of the sample selection terms in the two equations. At least one sample selectivity term is significant in each of the two earnings equations. Since $\lambda_{41} > 0$ and $\lambda_{42} > 0$, the negative coefficient estimated for λ_{41} implies a negative selection into the migrant group in terms of unobserved characteristics. The coefficient of λ_{42}, on the other hand, is insignificant and close to zero in magnitude. These imply that, immigrants' earnings in Canada are less, ceteris paribus, than what the earnings of non-migrants would have been in Canada had they migrated. That is, if host country is trying to select individuals with higher motivation, adaptability etc. under personal suitability factor, these individuals are less likely to apply for migration. These results and the previous ones related to schooling imply a negative self-selection at the application stage in terms of both observed and unobserved characteristics. For the source country earnings equation the coefficient estimates for the sample selectivity terms are negative. Given that $\lambda_1^* < 0$ and $\lambda_2^* < 0$ these coefficient estimates suggest a

positive selection among the non-migrants. That is, non-migrants earn more, ceteris paribus, than what the migrants in the source would earn had they not migrated to the host country.

In tables 4 and 5 the parameter estimate related to λ_{41} is much larger than that for λ_{42}, and the parameter estimate related to λ_1^* is much larger than that for λ_2^*. The larger coefficients for λ_{41} and λ_1^* correspond to the self-selection among the individuals at the application stage whereas the other two refer to the selection by the host country during the review process. This result shows that selectivity in terms of unobserved skills among the migrants is mostly due to the self-selection that occurs at the application stage which corresponds to the individuals' assessment of returns to the migration. While the host country has effective means of selecting individuals in terms of easily observed characteristics, such as education, it is much harder to select individuals based on other characteristics such as motivation.

The results presented here are also consistent with the Roy model. Given that the U.S. income distribution has higher income inequality than Canada and earnings between the two countries are positively and strongly correlated, the Roy model would predict that unobserved skills of applicants for migration from US to Canada will be below-average. The results suggesting a negative selection among applicants and a positive selection among non-migrants provide evidence supporting the predictions of the Roy model.

Using the parameter estimates of the earnings equations, earnings in the U.S. can be predicted for both immigrants and non-migrants. Figure 1 presents the distribution of predicted earnings in the U.S. for the migrants and figure 2 presents the same for non-migrants. This shows from which part of the U.S. income distribution the immigrants are coming. Comparing the two figures it becomes evident that almost half of the immigrants are coming from the upper end of the US income distribution corresponding to 90^{th} percentile or higher. While the previous discussions showed that immigrants are negatively self-selected in terms of observed and

unobserved characteristics, the review step by the host country tries to select those with higher observed and unobserved skills. The relative position of immigrants in the U.S. income distribution shows that self-selection of immigrants and their selection by the host country may move in the opposite directions. As in the case of observed characteristics the selection by the host country may outweigh the first one. This may be due to availability of effective policy instruments such as the points system used in the Canadian context for selecting immigrants. For unobserved characteristics, however, a positive selection is much harder to achieve by the host country. The results suggest that review step can not overturn the negative self-selection occurring at the application stage in terms of unobserved skills.

4.3 Results from the Structural Bivariate Probit Model

Estimation of the structural bivariate probit model given by (11.3) is carried out as outlined above. The model discussed in section 2 states that individuals will apply for migration if permanent income in the host country is greater than that in the source country net of moving costs. The use of predicted earnings differentials $\left\{(\hat{X_{1i}}\beta_1)-(\hat{X_{2i}}\beta_2)\right\}$ as a measure of permanent income differential between two countries deserves some comment. Ideally present value of life time earnings would be calculated for any given individual in both the source and host countries. However, this would require a much richer data than what is used in this study. For example, for the immigrants the data provides up to eight observations for a given individual and estimating life time earnings would necessarily require an assumption of no cohort effects and in some cases predictions out of the sample. Using estimated earnings differential has its problems as well. It is well known that immigrant may experience an assimilation process with low earnings in first few years after migration and higher earnings in later years. Therefore, at certain points of

the age-earnings or experience-earnings profiles although individuals may not have big gains or even realize some losses, this may be compensated in the years ahead as years of residence in the host country increases. Using the predicted earnings differentials $\left\{(\hat{X_{1i}}\beta_1)-(\hat{X_{2i}}\beta_2)\right\}$ for a given individual at the given level of experience and schooling in the data gives the difference between the experience-earnings profiles between the two countries at a single point. The coefficient estimate on this differential then tests whether individuals become more likely to apply for migration as this differential increases, providing evidence on the hypothesis that individuals choose to live in countries that provide higher permanent incomes. An alternative approach is adopted by Robinson and Tomes (1980) who partition their sample according to the experience and estimate their model for each experience interval. Then, the expectation is that the predictions of the model will be observed for at least one experience group. Given the small sample size of migrants in my data I am prohibited from adopting this approach. In the estimation of earnings differentials marital status is also taken as a permanent characteristic while individuals may change their marital status. Another important issue is that tax rates on a given level of income are different across the two countries. However, taking into account the tax rates also requires an assessment of the differences in the provision of public goods in the two countries. While in general taxes are higher in Canada, some goods such as health care, education are mostly provided by the government. The lower cost of such goods will be assumed to compensate for higher taxes.

In estimation of the structural probit, following Robinson and Tomes (1980), age and its square is included as a partial control for the fact that wage gain is estimated at a point in the life cycle rather than using the difference in permanent income streams. The moving costs $c_{mi} = X_{3i}\beta_3 + \varepsilon_{3i}$ has the same specification as discussed before where

$$X_{3i} = \{\text{family size, marital status, schooling}\}_i$$

While c_{mi} includes family size that increases moving costs due to change of schools for kids etc., it doesn't appear in the earnings equation. Similarly experience is assumed to affect migration through its effect on wages, however, it doesn't affect moving costs. These exclusion restrictions allow identification of parameters in structural index I_1. Table 6 presents the results from the estimation of the structural model. The standard errors are subject to a downward bias since estimates of the wage differences rather than the true (unknown) values are being used. In a similar setting Robinson and Tomes (1980) report the likely magnitude of this bias in a modified version of their model to be 25 per cent. The discussions below are subject to this caveat.

The important result from estimation of the structural model in table 6 is the coefficient for the earnings differential which is positive and significant. Subject to the caveat concerning significance this provides evidence in favor of the model[18]. This result becomes more impressive when viewed taking into account the fact that the model is tested for all experience groups combined together and also against various data issues arising from use of data from two different countries. Although much of the immigration literature studying international migration relies on the hypothesis that individuals respond to the earnings differentials across countries, to my knowledge, this result provides the first direct evidence supporting this hypothesis.

Other results in the first and second selection index are similar to the reduced form results. In the first selection index that refers to the application stage age variables are significant suggesting that marginal propensity to migrate decreases with age. A difference from the reduced form results is the positive and significant coefficient on years of schooling. Controlling for the earnings differential higher education increases mobility which may be due to better information

[18] Even if the correct standard error was twice the size of the reported standard error in the table, the estimated coefficient for the earnings differential would still be significant at the 1 % level.

or reduced costs of moving. Family size, which enters with a negative coefficient, is an important deterrent to migration whereas married variable enters with a positive coefficient increasing mobility. Single individuals who may have stronger connections to their families (e.g. because of living with parents) may be more tied stayers than married people. In the second selection index that refers to the probability of acceptance, results similar to the reduced form specification are obtained. Higher education and being in a white collar or professional occupation increase the probability of acceptance while increasing age decreases this probability consistent with the structure of the point system.

5 Conclusions

This paper specified and estimated a structural model of international migration using micro data. This provides a direct test of human capital theory that suggests that individuals respond to the earnings differentials across countries while making their migration decisions. The model also recognizes that there are two decision makers affecting the outcome of international migration, i.e. the individual and the host country. Existing studies of international migration abstract from this and estimate the outcome of migration without being able to separately identify the factors determining the decision of these two players.

The model proposed by Borjas(1987) based on Roy has predictions for the selection of observed and unobserved characteristics. It is harder to test the predictions of this model for the unobserved characteristics and there are no direct tests available in the literature. The indirect tests employed in the literature by Borjas and other researchers provided mixed evidence and also stirred a lot of discussion on whether there is favorable selection amongst immigrants. The

framework of this paper which also takes into account the selection by the host country provides the first direct evidence on the issue.

The empirical results in this paper support the human capital model with the evidence showing that wage differentials between the host and the source country affect the migration decision in the predicted direction. The estimates of life time earnings could be improved by using panel data for both migrants and non-migrants and taking into account individual heterogeneity. While incorporating these refinements is important in this setting, the Census data used for non-migrants doesn't allow pursuing this avenue. The results also show that both the host country and the individual each have significant impact on the outcome of international migration. As shown by the reduced form results, while schooling increase the probability of acceptance by the host country it decreases the probability of application due to the higher returns to education in US. Finally, estimated wage equations and predicted earnings of immigrants in the source country provide evidence on selection issues. The results suggest a negative self-selection at the application stage both in terms of observed and unobserved characteristics which is consistent with the predictions of the Roy model. However, when the predicted earnings of immigrants in the source country is compared to the predicted earnings of non-migrants it is shown that they come from the upper end of the income distribution. This result is due to positive selection that occurs at the review step by the host country which shows the importance of rationing of visas by the source country in determining the resulting selectivity of immigrants.

Appendix A - Immigration Policy of Canada

Canada admits immigrants under four main categories: 1) Independent, 2) The refugee and humanitarian class, 3) "Sponsored" dependents (husband, wife, fiancée, generally close relatives), 4) "Nominated relative" (apply likewise to close relatives).

"Nominated relatives" and "Independent applicants" constituted about 50% of the total number of migrants admitted in 1990 with over half of them being migrants under Independent Class. The applications under "independent" and the "nominated relative" classes are subject to the point system where the pass mark was 70 points over the period we study.

The following table outlines the point system that was effective over 1986-1990:

Point System 1986-1990

Category	Potential Points
Education	12
Special Vocational Preparation	15
Experience	8
Occupational Demand	10
Arranged Employment/Designated Occ.	10
Age	10
Knowledge of French and English	15
Personal Suitability	10
Demographic Factor	10
Total	105[19]

The number of points to be awarded for each factor in the point system, except the personal suitability, is determined according to a schedule. Occupations are listed in 7-digit detail with each occupation on the list assigned a maximum of 10 points based on the deemed demand for that occupation in the Canadian labour market. Specific vocational preparation (SVP) points

[19] With the demographic factor set at 5, the maximum points possible from the characteristics listed is equal to 95. If the immigration officer decides to use his discretionary power he can award up to 10 points. This is different than the personal suitability. Use of discretion is very rare and nobody in our sample passed with discretion.

depend on the amount of formal training required for average performance in that occupation. Applicants with higher levels of education and experience get higher points in each of these categories, while the number of points for age category goes down with increasing age. Those who have a job offer validated by Human Resources and Development Canada get extra 10 points. Ability to read, write and speak in the two official languages determines the number of points under the language factor. Demographic factor (or levels), on the other hand, is set at 5 for everyone over the period we study. By setting different levels it is aimed to control the number of people entering over a period.

The only subjective factor in the point system is the personal suitability, for which a maximum of ten points is available. The visa officer will make arrangements for an interview with the applicant where he considers the latter could meet the selection criteria, based on the information provided and the points the applicant could be awarded for personal suitability. According to the immigration manual determination of the number of units of assessment to be awarded to an applicant rests on the judgment of the interviewing officer. The qualities of adaptability, motivation, initiative, resourcefulness, and such other attributes admirable or otherwise, as the applicant may display, are characteristics on which the officer may base his determination

Table 1
Definitions of Variables

Variable Type	Definition	Mnemonic
Migration	1: if application for migration received between 1986-1990	migr
	And accepted; 0 otherwise	
Schooling	Years of schooling completed	nyrssch
	(Years of schooling completed)/10	yrssch
Age	(Age)/10 and its square	age
		agesq
Experience	Labour market experience= (Age-Yrssch-6); and its square	exp
		expsq
Marital Status	Married: 1: if married or common law; 0: otherwise (Reference group is single-never married, widow or separated)	married
Dependents	Dep1: 1:if number of dependents is equal to 1; 0: otherwise	dep1
	Dep2: 1:if number of dependents is equal to 2; 0: otherwise	dep2
	Dep3: 1:if number of dependents is equal to 3; 0: otherwise	dep3
	Depg3: 1:if number of dependents is greater than 3; 0: otherwise (Reference group is those with no dependents – see notes below for details)	depg3
Occupation	Exec: 1: if executive, administrative or managerial occup.; 0: otherwise	exec
	Prof: 1: if professional specialty occup.; 0: otherwise	prof
	Techn: 1: if technicians and related occup.; 0: otherwise	techn
	Sales: 1: if sales occup.; 0: otherwise	sales
	Service: 1: if service occup.; 0: otherwise	service
	Farm: 1: if farming, forestry, and fishing occup.; 0: otherwise (Reference group is other blue-collar occupations)	farm

Notes for Table 1:
Number of dependents is calculated as follows:

- Migrants: number of potential visas to be issued to a principal applicant and the dependents excluding the principal applicant and the spouse. Visas are issued to spouse, children who are never married under the age 18, or children over 18 who are dependent on their parents due to for example a long-term illness.
- Non-migrants: number of persons in the family excluding the parents.

Table 2
Descriptive Statistics

	Migrants		Non-Migrants	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**
age	3.79	0.81	4.02	1.25
yrssch	1.77	0.33	1.30	0.29
dep1	0.15	0.36	0.21	0.40
dep2	0.20	0.40	0.23	0.42
dep3	0.08	0.27	0.10	0.30
depg3	0.01	0.10	0.08	0.27
married	0.74	0.43	0.66	0.47
exec	0.35	0.47	0.18	0.38
prof	0.54	0.49	0.10	0.30
farm	0.004	0.06	0.040	0.19
sales	0.014	0.11	0.151	0.35
service	0.010	0.09	0.078	0.26
techn	0.018	0.13	0.034	0.18
N	2,500		10,305	

Note: For dummy variables, mean values reported correspond to the proportion of individuals in the sample with that characteristic.

Table 3
Reduced Form Two-Step Model

		Coefficient	Std. Error
First Selection Index (Application Decision)	constant	-3.28	(0.36)
	age	3.16	(0.21)
	agesq	-0.37	(0.02)
	yrssch	-0.41	(0.15)
	dep1	-1.03	(0.08)
	dep2	-1.11	(0.08)
	dep3	-1.07	(0.09)
	depg3	-2.04	(0.13)
	married	0.58	(0.05)
		Coefficient	Std. Error
Second Selection Index (Application Review)	constant	-4.94	(0.07)
	age	-0.15	(0.01)
	yrssch	1.06	(0.03)
	exec	0.36	(0.01)
	prof	0.34	(0.01)
	techn	-0.004	(0.02)
	sales	-0.26	(0.02)
	service	-0.07	(0.03)
	farm	-0.06	(0.04)
	rho($\varepsilon, \varepsilon_4$)	-0.56	(0.01)
	N	12,805	

Note: Estimation results from a bivariate probit model with partial observability: Standard errors are given in parentheses.

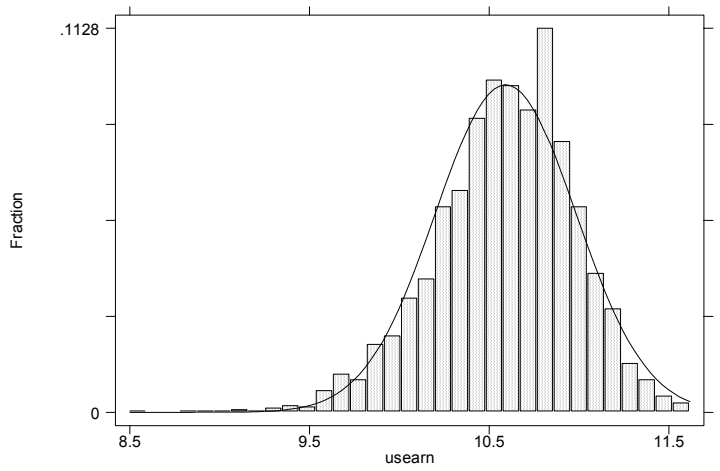
Table 4
Coefficient Estimates - Host Country Earnings Equation
Corrected for Sample Selectivity

Variable	Coefficient	Std. Error
constant	9.031	(.096)
nYrssch	0.051	(.004)
exp	0.048	(.005)
expsq	-0.0008	(.0001)
married	0.319	(.024)
λ_{41}	-0.129	(.026)
λ_{42}	0.0003	(.007)
N	10472	

Table 5
Coefficient Estimates - Source Country Earnings Equation
Corrected for Sample Selectivity

Variable	Coefficient	Std. Error
constant	7.953	(.023)
nYrssch	0.098	(.017)
exp	0.067	(.001)
expsq	-0.001	(.00002)
married	0.386	(.008)
λ_1^*	-0.091	(.009)
λ_2^*	-0.004	(.001)
N	54386	

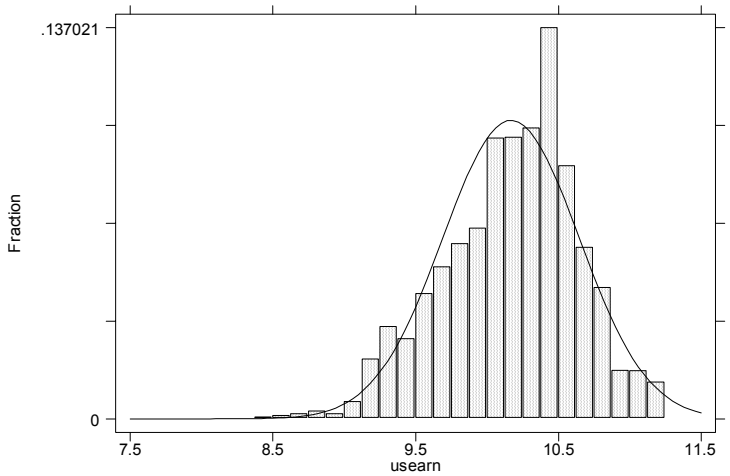
Figure 1 - Predicted Log Earnings in US for Immigrants



Descriptive Statistics for Figure 1

		Percentiles								
N	mean	1%	5%	10%	25%	50%	75%	90%	95%	99%
2500	10.59	9.62	9.88	10.07	10.36	10.62	10.85	11.06	11.19	11.42

Figure 2 - Predicted Log Earnings in US for Non-migrants



Descriptive Statistics for Figure 2

		Percentiles								
N	mean	1%	5%	10%	25%	50%	75%	90%	95%	99%
10305	10.17	9.02	9.32	9.51	9.87	10.23	10.47	10.72	10.84	11.17

Table 6
Structural Two-Step Model

		Coefficient	Std. Error
First Selection Index (Application Decision)	dlnw	5.66	(1.13)
	age	2.57	(0.49)
	agesq	-0.21	(0.08)
	yrssch	2.86	(0.54)
	dep1	-0.22	(0.05)
	dep2	-0.24	(0.05)
	dep3	-0.21	(0.05)
	depg3	-0.61	(0.11)
	married	0.49	(0.09)
	constant	-14.44	(1.92)
		Coefficient	Std. Error
Second Selection Index (Application Review)	age	-1.08	(0.19)
	yrssch	0.87	(0.13)
	exec	0.51	(0.08)
	prof	0.45	(0.07)
	techn	-0.1	(0.09)
	sales	-0.36	(0.08)
	service	-0.13	(0.08)
	farm	0.005	(0.07)
	constant	1.3	(1.6)
	rho($\varepsilon, \varepsilon_4$)	-0.44	(0.21)
	N	12,805	

Note: Estimation results from a bivariate probit model with partial observability: Standard errors are given in parentheses.

Bibliography

Adams, R. H. Jr. (1993). The Economic and Demographic Determinants of International Migration in Rural Egypt. *The Journal of Development Studies*, 30(1), 146-167.

Amemiya, T. (1978). The Estimation of a Simultaneous Equation Generalized Probit Model. *Econometrica*, 46, 1193-1205.

Aydemir, A. B. (2003). Estimating the Effects of Immigration Policy: Disentangling Policy Effects from Self-Selection, mimeo, University of Western Ontario, Department of Economics.

Borjas, G. J. (1987). Self-Selection and the Earnings of Immigrants. *AER*, 77(4), 531-551.

Borjas, G. J. (1990). Self-Selection and the Earnings of Immigrants: Reply. *AER*, 80(1), 305-308.

Borjas, G. J. (1991). Immigration and Self-Selection. In *Immigration, Trade and the Labor Market*, eds. R. B. Freeman and J. Abowd., 29-76. Chicago: U. of Chicago Press.

Carpentier, Alain and Gerard Pinsonneault (1994). *Representativeness Study of Immigrants Included in the Immigrant Data Bank*. Studies and Research Branch, Ministere des Affairs Internationales, de l'Immigration et des Communautes Culturelles, Quebec.

Chiswick, B. R. (1978). The Effect of Americanization on the Earnings of Foreign-Born Men. *Journal of Political Economy*, October 1978, 86, 897-921.

Chiswick, B. R. (1999). Are Immigrants Favorably Self-Selected?. *AEA Papers and Proceedings*, 89(2), 181-185.

Chiswick, B. R. (2000). Are Immigrants Favorably Self-Selected An Economic Analysis? In *Migration Theory: Talking Across the Disciplines*, eds. C. D. Brettell and J. F. Hollofield, New York: Routledge Press.

Cobb-Clark, Deborah A. (1993). Immigrant Selectivity and Wages: The Evidence for Women. *AER*, 83(4), 986-993.

Greenwood, M. J. and J. M. McDowell (1991). Differential Economic Opportunity, Transferability of Skills, and Immigration to the United States and Canada. *The Review of Economics and Statistics*, 73(4), 612-623.

Heckman, J. J. (1976). Simultaneous Equation Models with Continuous and Discrete Endogenous Variables and Structural Shifts. In S. M. Goldfeld and R. E. Quandt, eds., *Studies in Nonlinear Estimation*, 235-272. Ballinger, Cambridge, MA.

Heckman, J. J. (1978). Dummy Endogenous Variables in a Simultaneous Equation System. *Econometrica*, 46, 931-959.

Jasso G. and M. R. Rosenzweig (1990). Self-Selection and the Earnings of Immigrants: Comment. *AER*, 80(1), 298-304.

Lucas, R. E. (1985). Migration Amongst the Botswana. *The Economic Journal*, 95, 358-382.

Manski, C. F. and S. R. Lerman (1977). The Estimation of Choice Probabilities from Choice Based Samples. *Econometrica*, 45(8), 1977-88.

Mincer, J. (1974). *Schooling Experience and Earnings*. NBER, Columbia University Press, New York.

Mincer, J. (1978). Family Migration Decision. *JPE*, 86, 749-773.

O Grada (1986). Determinants of Irish Emigration: A Note. *International Migration Review*, 20(3), 651-656.

Poirier, D. J. (1980). Partial Observability in Bivariate Probit Models. *Journal of Econometrics*, 12, 209-217.

Robinson, C. and N. Tomes (1982). Self Selection and Interprovincial Migration in Canada. *CJE*, 15(3), 474-502.

Roy, A. D. (1951). Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers*, 3, 135-146.

Stark, O. and J. E. Taylor (1991), Migration Incentives, Migration Types: The role of Relative Deprivation. *The Economic Journal*, vol. 101, pp. 1163-1178.

Taylor, J. E. (1986). Differential Migration, Networks, Information and Risk. In: *Research in Human Capital and Development: Migration, Human Capital and Development*. 4, 147-171.

Tunali, I. (1986). A General Structure for Models of Double-Selection and an Application to a Joint Migration/Earnings Process with Remigration. In: Ronald E. Ehrenberg, ed., *Research in Labor Economics*, 235-282, 8, Part B.