DISCUSSION PAPER SERIES

Identification and Estimation of Economic Models of Outmigration Using Panel Attrition

Charles Bellemare

March 2004

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Identification and Estimation of Economic Models of Outmigration Using Panel Attrition

Charles Bellemare
*CentER, Tilburg University
and IZA Bonn*

ABSTRACT

Identification and Estimation of Economic Models of Outmigration Using Panel Attrition[*]

Because their departures are difficultly observed, little is known about the performance of immigrants who leave a region and move to another. This paper shows conditions under which the (conditional) outmigration probability, work probability and the expected earnings of outmigrants are nonparametrically identified using data on immigrant sample attrition. We present an econometric model which extracts the information on outmigration behavior from sample attrition and allows to incorporate unobserved heterogeneity in the choice process in a very natural way, a feature not previously possible. The method presented in the paper is general enough to estimate models of internal and external migration. We apply our framework to estimate a panel data model where immigrants can choose whether to stay or leave a country. The model is estimated using the German Socio-Economic Panel (GSOEP). Outmigrants are found to have significantly lower labor market earnings and work propensities than permanent migrants.

JEL Classification: J24, J61

Keywords: outmigration, identification, panel data models

Charles Bellemare
Department of Econometrics
Tilburg University
Warandelaan 2
5000 LE Tilburg
The Netherlands
Email: c.bellemare@uvt.nl

1 Introduction

The widespread and often sizeable flows of immigrants leaving their adoptive home-lands, which we will refer to as outmigration, have recently received a lot of attention. The case of Germany is a particularly revealing example, with an estimated yearly out-flow of half a million immigrants over the last decade (OECD, 2001). Several theories have been put forward to motivate outmigration movements. Theories based on earnings differentials between the current and new destination (Harris and Todaro, 1970), higher marginal utility of consumption in the home country (Djajic and Milbourne, 1988), high returns to human capital investments in the host country (Dustmann, 1993), informa-tion dissemination (Stark, 1995), credit market rationing in the native country (Mesnard, 2001), and several sociological factors such as family unification, health satisfaction, feel-ings of being integrated in society (Stark, 1998) , and the quality and productivity of an immigrant's social network (e.g. Carrington, Detragiache and Vishwanath (1996)).

These theories do not trivially predict a specific composition of departing immigrants. It could be the case for example that economically successful immigrants with a rela-tively higher marginal utility of consumption in their native country might opt to leave despite relatively lower earnings in that country, while persistently unsuccessful immi-grants might find a worthwhile to keep on searching for better labor market prospects and move to a new destination. The lack of a clear-cut theoretical prediction concerning the economic performance of outmigrants complicates the design of immigration poli-cies which are often tailored around the notion that labor market equilibriums may be affected by outmigration flows of non-randomly selected workers. Careful forecasting of the quality of the migration flows is thus necessary if immigration policies are to meet the future needs of the labor market. Additionally, recent theoretical and empirical evidence has suggested that measures of economic progress of immigrants in their host country can be adversely affected by selective outmigration (Schultz, 1998, and Edin, LaLonde, and Aslund, 2000). Policies aimed at improving the labor market integration of its immigrant population may thus also be misguided if it relies on these potentially biased measures of immigrant assimilation.

In this paper, we are interested in testing competing explanations of outmigration decisions while at the same time assessing the robustness of measures of economic assim-ilation to outmigration flows. We do so by estimating a limited-dependent variable panel data model where labor market earnings, work decisions, and outmigration decisions are jointly determined and depend on earnings differentials, family unification, credit market rationing, health satisfaction, and feelings of being integrated in the host society. Unob-servable characteristics such as an immigrant's ability or the quality and productivity of

his social network are incorporated in the model as individual specific unobserved heterogeneity components. The general error structure of the model allows to test whether outmigrants are self-selected in terms of labor market earnings and in terms of work status. The model also allows to assess the impact of outmigration selection on estimates of measures of economic assimilation rates. Compared to existing empirical models of outmigration (see below), our model has the advantage of both incorporating the decision to work in a natural way and characterizing the relationship between work status and outmigration.

Estimation of our model requires panel data on immigrants followed over a relatively long period of time. As Dustmann (2002) recently pointed out, interesting empirical analysis of outmigration decisions is limited by the fact that data sets rarely contain information on outmigration decisions[1]. Rather, they typically contain information on whether immigrants fail to be successfully reinterviewed in subsequent waves, which may or may not be due to selective outmigration. Existing empirical evidence on the economic performance of outmigrants is pretty much tied to the strategy used to identify the economic parameters characterizing the performance of outmigrants without observing outmigration decisions. Jasso and Rosenzweig (1990) identify the direction of outmigration earnings selectivity by comparing the skill composition of specific cohorts over time. Hu (2000) and Lubotsky (2000) estimate the parameters of the earnings function of immigrants who remain in the country, controlling for non-random outmigration selectivity by matching cross-section data sets and longitudinal social security earnings records. These approaches are not without their own limitations-they do not identify the level of earnings of outmigrants, census and earnings records often have little information on both the human capital level and sociological characteristics of migrants which are required to test outmigration theories [2], they do not easily allow migration duration decisions to depend on unobserved characteristics such as inherent ability or the quality of an immigrant's social network which, at least on theoretical grounds, are important determinants of migration durations.

In this paper, we present conditions under which the outmigration probability, the work probability, and the expected labor market earnings of outmigrants are, conditionally on observable characteristics, nonparametrically identified from survey data with sample attrition. The cornerstone of our identification approach consists of using panel attrition as a proxy variable for outmigration, and subsequently separating attrition which is not due to outmigration movements from real outmigration decisions. Our approach overcomes several of the shortcomings of earlier approaches used to identify the economic performance of outmigrants. First, our approach is to our knowledge the first one

which has the potential to provide nonparametric identification of the economic performance and movements of outmigrants. Second, because the approach proposed uses survey panel data instead of census data, unobserved heterogeneity can easily be introduced in the model. Third, the approach is general enough to be applied to any country with an ongoing panel of immigrants (examples of countries with such panels are Canada, Mexico, Germany, and the United-States.) and can be easily extended to estimate many different types of economic models of outmigration. An example of this flexibility is our ability to analyze the interaction between earnings, work, and outmigration decisions in a unified framework which was not previously possible without observing actual migration decisions. Finally, the estimator proposed is simple to apply, and is similar in spirit to estimators proposed to deal with misclassification of a discrete dependent variable (e.g. Hausmann, Abrevaya, and Scott-Morton, 1998).

We estimate our model using the public use file of the GSOEP. We use data on native Germans as a reference group to compute earnings assimilation rates for our immigrant sample. We show how the availability of a sample drawn from the native population has the additional advantage of providing a natural way to test some of the identifying assumptions of the model. We estimate the annual outmigration rate in our sample to be roughly 3% per year over our time horizon. Our simulation results indicate that average log earnings of outmigrants remained roughly 18% lower than those of immigrant stayers, a clear indication that outmigrants are drawn from the bottom of the income distribution. Moreover, outmigrants are shown to have work probability 25% to 45% lower than that of immigrant stayers over the period considered. Finally, we do not find that assimilation rates are particularly sensitive to outmigration, which contrasts with existing results found in the literature.

The rest of this paper is organized as follows. Section 2 presents our approach to identify the economic parameters of interest. Section 3 presents the econometric model used to model outmigration in conjunction with the work decision and labor market earnings. Section 4 presents the data used in the paper. Section 5 discusses the empirical results of the model and tests for the presence of outmigration bias. It further presents some simulation results used to evaluate the fit of the model and to quantify the economic performance of outmigrants. Section 6 concludes.

2 Identification of outmigration parameters

Each immigrant of a population living in the host country is characterized in a given time period by the vector (w, p, r^u, x, z, s) where w denotes his potential labor market earn-

ings conditional on characteristics x, p is a binary variable taking a value of 1 when the immigrant works and whose outcome is conditioned on a vector of observable characteristics z, r^u is a binary indicator taking a value of 1 when the immigrant outmigrates in the next time period and we condition this outcome on a vector of characteristics s. We define X as the vector of all distinct elements of (x, z, s). We are interested in making inferences on $\Pr(p = 1|r^u = 1, X)$, conditional work probability of an outmigrant, and on $E\{w|p = 1, r^u = 1, X\}$, the conditional expected earnings of outmigrants. The inferential problem consist of identifying these quantities when, instead of observing outmigration, we observe a proxy variable r^o, panel attrition, which takes a value of 1 when the immigrant leaves the panel in the following period. Outmigration and attrition are related because an immigrant who leaves the country must also leave the panel with probability 1.[3]

We illustrate the identification problems for the case where we want to make inferences on $E\{w|p = 1, r^u = 1, X\}$. However, all results extend directly to identification of $\Pr(p = 1|r^u = 1, X)$.[4] The conditional expected labor market earnings of immigrants who leave the panel in the next time period can be expressed, using iterated expectations, as

$$
\begin{aligned}
E\{w|p = 1, r^o = 1, X\} &= E\{w|p = 1, r^o = 1, r^u = 1, X\} \cdot \Pr(r^u = 1|r^o = 1, X) \\
&\quad + E\{w|p = 1, r^o = 1, r^u = 0, X\} \cdot \Pr(r^u = 0|r^o = 1, X) \\
&= E\{w|p = 1, r^u = 1, X\} \cdot \Pr(r^u = 1|r^o = 1, X) \\
&\quad + E\{w|p = 1, r^u = 0, X\} \cdot \Pr(r^u = 0|r^o = 1, X) \quad (1)
\end{aligned}
$$

where the second equality follows from the fact that once we know if an immigrant left or not the country, observing whether he left of not the panel does not contain any additional information on the earnings of this immigrant conditional on X.[5] An immediate consequence of equation (1) is that using panel attrition as a proxy variable for outmigration in itself will give biased and inconsistent estimates of the conditional earnings of outmigrants. This is so because the conditional expected earnings of immigrants who leave the panel will in general be a weighted average of the conditional expected earnings of outmigrants mixed with the conditional earnings of immigrants who remain in the host country. The mixing probabilities control the size of the bias. The key parameter is $\Pr(r^u = 0|r^o = 1, X)$, which represents the probability that an immigrant stays in the the host country given that he is observed to leave the panel. The higher this probability, the higher will be the bias. If every immigrant who leaves the panel also leaves the country, r^o would perfectly measure outmigration, $\Pr(r^u = 0|r^o = 1, X)$ would be equal to zero, and the bias would be zero.

Next, we follow the same steps to derive the conditional expected earnings of immi-

grants who remain in the panel

$$E\{w|p=1,r^o=0,X\} = E\{w|p=1,r^u=1,X\}\cdot\Pr(r^u=1|r^o=0,X) \quad (2)$$
$$+E\{w|p=1,r^u=0,X\}\cdot\Pr(r^u=0|r^o=0,X)$$

Because an immigrant cannot be observed to have left the country given he is observed to be in the panel, $\Pr(r^u=1|r^o=0,X)=0$, $\Pr(r^u=0|r^o=0,X)=1$ and (2) simplifies to

$$E\{w|p=1,r^o=0,X\}=E\{w|p=1,r^u=0,X\} \quad (3)$$

which indicates that the conditional earnings of immigrants who remain in the panel co-incides with the expected earnings of immigrants who remain in the host country, and it follows that $E\{w|p=1,r^u=0,X\}$ is nonparametrically identified from data on immigrants in the sample. Substituting (3) in (1) we obtain

$$E\{w|p=1,r^o=1,X\} = E\{w|p=1,r^u=1,X\}\cdot\Pr(r^u=1|r^o=1,X)$$
$$+E\{w|p=1,r^o=0,X\}\cdot\Pr(r^u=0|r^o=1,X)$$

which can be solved in terms of $E\{w|p=1,r^u=1,X\}$, the parameter we hope to identify,

$$E\{w|p=1,r^u=1,X\} = E\{w|p=1,r^o=1,X\}\cdot W_1(X)^{-1}$$
$$-E\{w|p=1,r^o=0,X\}\cdot W_0(X)W_1(X)^{-1} \quad (4)$$

Equation (4) shows that the conditional expected earnings of outmigrants can be ex-pressed as a weighted average of two conditional expectations which are identified from the data. If the weights can be identified, then the conditional earnings of outmigrants will be identified. Applying Baye's rule, the weights are given by

$$W_0(X) = \Pr(r^u=0|r^o=1,X)$$
$$= \Pr(r^o=1|r^u=0,X)\frac{\Pr(r^u=0|X)}{\Pr(r^o=1|X)}$$
$$W_1(X) = \Pr(r^u=1|r^o=1,X)$$
$$= \underbrace{\Pr(r^o=1|r^u=1,X)}_{=1}\frac{\Pr(r^u=1|X)}{\Pr(r^o=1|X)}$$
$$= \frac{\Pr(r^u=1|X)}{\Pr(r^o=1|X)}$$

$\Pr(r^o=1|X)$ is identified from the attrition data. What remains to be identified is $\Pr(r^u|X)$ and $\Pr(r^o=1|r^u=0,X)$. It is clear assumptions must be placed on the data generat-ing process to identify the relationship between the observable attrition indicator r^o and the unobservable outmigration indicator r^u. To simplify the notation, we will denote

$\alpha_{10}(X) \equiv \Pr(r^o = 1 | r^u = 0, X)$. Using the law of total probability, the attrition probability can be expressed in general terms as

$$\Pr(r^o = 1 | X) = \alpha_{10}(X) + [1 - \alpha_{10}(X)] \cdot \Pr(r^u = 1 | X) \tag{5}$$

and the probability of remaining in the sample as

$$\Pr(r^o = 0 | X) = 1 - \{\alpha_{10}(X) + [1 - \alpha_{10}(X)] \cdot \Pr(r^u = 1 | X)\} \tag{6}$$

Several assumptions can be imposed on the data generating process in order to identify both $\alpha_{10}(X)$ and $\Pr(r^u = 1 | X)$.

Assumption A1 $\alpha_{10}(X) = \alpha_{10} \geq 0$

Proposition 1 *If A1 holds and there exists a X such that $\Pr(r^u = 1 | X) = 0$, $E\{w | p = 1, r^u = 1, X\}$ is nonparametrically identified.*

Proof. Given the foregoing discussion, it suffices to establish that the conditions of the Proposition identify the weights $W_0(X)$ and $W_1(X)$. From (5), it follows that $\Pr(r^o = 1 | X) = \alpha_{10}$ which identifies α_{10} from limit observations satisfying $\Pr(r^u = 1 | X) = 0$. Given α_{10} is identified, $\Pr(r^u = 1 | X)$ is identified from (6), which implies that the weights $W_0(X)$ and $W_0(X)$ are both nonparametrically identified. ∎

Proposition 1 shows that if attrition which is not due to outmigration is random in the population (A1), all important economic parameters characterizing outmigration behavior can be recovered from the data if there exists a sample of permanent migrants; i.e. immigrants whose outmigration probability is close to zero. In practice, this does not seem to be a strong requirement, especially for western countries where permanent migration is known to occur at a very large scale (OECD, 2001). Note that in practice, A1 needs not to hold if attrition which is not related to outmigration does not vary in the population (i.e. if the variance $V(\alpha_{10}(X)) \approx 0$). This can be verified for example by computing marginal effects from binary choice regressions on attrition outcomes for a sample of individuals who by construction do not outmigrate, and test if these effects are small. Natives living in the host country is one example of a sample not prone to outmigration.

If attrition which is not due to outmigration is believed to be related to observable factors which induce sufficient variation in the attrition process across individuals, nonparametric identification of the economic parameters of outmigrants requires some exclusion restriction.

Assumption A2 $\alpha_{10}(X) = \alpha_{10}(X_1) \geq 0$ where $X = (X_1', X_2')'$

Proposition 2 *If A2 holds and there exists a X_2^* given X_1 such that $\Pr\left(r^u = 1|X_1, X_2^*\right) = 0$,*

$E\left\{w|p = 1, r^u = 1, X\right\}$ is nonparametrically identified

Proof. For a given $X = (X_1', X_2')'$, $\alpha_{10}(X) = \alpha_{10}(X_1)$ from A2. Using $\Pr\left(r^u = 1|X_1, X_2^*\right) = 0$, the subsample of immigrants $(X_1', X_2^{*\prime})'$ identifies $\alpha_{10}(X_1)$ from (5). Given $\alpha_{10}(X_1)$ is identified, $\Pr\left(r^u = 1|X\right)$ is identified from (6), which implies that the weights $W_0(X)$ and $W_0(X)$ are both nonparametrically identified. ∎

Partly because attrition and outmigration are very different processes, finding realistic exclusion restrictions which satisfy the requirements of Proposition 2 is not very restrictive. Usually outmigration is modelled as a life-cycle event, influenced by poor labor market performance, integration feelings, credit rationing in the home country and age at immigration. Whether attrition which is not due to outmigration is related to all these factors seems a priori unlikely, given that part of the survey non-response is generally based on respondents refusal to continue working with the survey agencies.

3   Parametric model and estimation method

In this section, we develop and estimate a parametric model which allows us to extract outmigration behavior from panel attrition. The choice of a parametric model is motivated by our desire to model selection into work and outmigration as a decision based on individual specific unobserved heterogeneity. We are not aware of any existing semiparametric technique which would allow us to estimate the system of equations presented below.

We have a measure of N immigrants in period 1, where immigrant i remains in the panel for T_i periods. For each immigrant i, we observe in period t, whether he works p_{it}, his monthly labor market earnings $e^{(w_{it})}$, and his attrition status r_{it}^o in the next period. The log of the potential labor market earnings is assumed to be generated by a log linear earnings equation

$$w_{it} = x_{it}'\beta + \eta_i^1 + \varepsilon_{it}^1 \tag{7}$$

where β are unknown parameters, η_i^1 is an unobserved time invariant individual specific component of income while ε_{it}^1 represents a stochastic shock. These labor market earnings are only observable when an immigrant works. The work decision p_{it} is assumed to be generated by a latent process

$$p_{it}^* = z_{it}'\theta + \eta_i^2 + \varepsilon_{it}^2 \tag{8}$$

8

where $\boldsymbol{\theta}$ are unknown parameters, η_i^2 is an unobserved component of work and ε_{it}^2 represents some stochastic shock to the work propensity. Participation is determined by the observation rule $p_{it} = 1\left[p_{it}^* > 0\right]$. When $p_{it} = 1$, earnings w_{it} are observed. Both η_i^1 and η_i^2 can be thought of capturing immigrants unobserved ability to generate higher earnings and to find jobs. They can also be thought of as including unobserved family background characteristics and preferences for work and leisure. Finally, an immigrant's unobservable outmigration propensity r_{it}^* is assumed to be determined by another latent process

$$r_{it}^* = s_{it}' \gamma + \eta_i^3 + \varepsilon_{it}^3 \tag{9}$$

where γ are unknown parameters, η_i^3 captures the individual specific attachment to his native country and ε_{it}^3 is a stochastic shock. The triplet $\left\{\eta_i^1, \eta_i^2.\eta_i^3\right\}$ is assumed to be observed by the immigrant who takes it into account when making his decisions but it is not observed by the econometrician. Let $r_{it}^u = 1\left[r_{it}^* > 0\right]$ be the decision rule governing the true outmigration decision in period $t + 1$. Outmigration r_{it}^u is unobservable. In our empirical application, we assume that A1 holds[6] and express the attrition probability as

$$\Pr\left(r_{it}^o = 1 | s_{it}\right) = \alpha_{10} + \left[1 - \alpha_{10}\right] \cdot \Pr\left(r_{it}^u = 1 | s_{it}\right) \tag{10}$$

Equation (10) is the sample counter part of equation (5).[7]

The earnings, work and outmigration outcomes are not likely to be independent of each other. This will not be independent if, for example, immigrants who find work very easily and/or who earn a high income are more reluctant to outmigrate. The unobserved heterogeneity components η_i^1, η_i^2 and η_i^3 can be treated either as fixed constants or as random variables. The main advantage of the fixed effect approach is that it does not require that included explanatory variables be strictly exogenous to the unobserved heterogeneity components $\left(\eta_i^1, \eta_i^2, \eta_i^3\right)$. However, estimation of fixed effects in nonlinear models remains today a sizeable complication, with very little guidance in the choice of models (see the recent review of Arellano and Honore, 2001). A second drawback of fixed effect estimation is that by treating the unobserved heterogeneity components as fixed, cross equation correlations which drive selection into work and outmigration based on unobservable individual characteristics are not identified. As the present paper is mainly concerned with selection issues, fixed effect estimation would limit our insights in the type of selections present in the data. We therefore introduce these dependencies by assuming that the stochastic time-invariant effects are independent and identically

normally distributed over time with mean 0 and covariance matrix

$$
\boldsymbol{\Omega} = \begin{bmatrix} \sigma_{\eta^1}^2 & \rho_{1,2}^{\eta}\sigma_{\eta^1}\sigma_{\eta^2} & \rho_{1,3}^{\eta}\sigma_{\eta^1}\sigma_{\eta^3} \\ \cdot & \sigma_{\eta^2}^2 & \rho_{2,3}^{\eta}\sigma_{\eta^2}\sigma_{\eta^3} \\ \cdot & \cdot & \sigma_{\eta^3}^2 \end{bmatrix}
$$

where $\sigma_{\eta^j}^2$ denotes the variances of the unobserved heterogeneity components, and $\rho_{i,j}^{\eta}$ denotes their correlations.[8] These correlations are indicative of whether or not immigrants self-select themselves into work and into outmigration based on their unobservable individual characteristics. A significant and positive $\rho_{1,2}^{\eta}$ indicates that individuals who are more likely to work are also more likely to have higher earnings, give observed characteristics. $\rho_{1,3}^{\eta}$ has a similar interpretation and is indicative of outmigration bias. This coefficient will be negative (positive) if immigrants who have a higher probability of outmigrating have below (above) average monthly earnings. Finally, $\rho_{2,3}^{\eta}$ can be interpreted as measuring outmigration bias in the work decision and whose sign has a similar interpretation.

Finally, we assume that the vector $\left[\varepsilon_{it}^1, \varepsilon_{it}^2, \varepsilon_{it}^3\right]'$ is i.i.d normally distributed with mean 0 and covariance matrix

$$
\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_w^2 & \rho_{1,2}^{\varepsilon}\sigma_w & \rho_{1,3}^{\varepsilon}\sigma_w \\ \cdot & 1 & \rho_{2,3}^{\varepsilon} \\ \cdot & \cdot & 1 \end{bmatrix}
$$

where σ_w^2 is the variance of log earnings, while the variances of the unobserved stochastic shocks entering the work and outmigration equations are set to 1 for identification purposes. Contemporaneous correlations between the three stochastic components are captured by the correlation coefficients $\rho_{1,2}^{\varepsilon}$ $\rho_{1,3}^{\varepsilon}$ and $\rho_{2,3}^{\varepsilon}$.

To simplify the presentation of the likelihood function, we divide the observable characteristics of immigrant i into a set $\boldsymbol{y}_i = \left\{p_{it}, r_{it}^o, w_{it} \cdot p_{it}\right\}_{t=1}^{T_i}$ of dependent variables, a set $\boldsymbol{X}_i = \left\{x_{it}, z_{it}, s_{it}\right\}_{t=1}^{T_i}$ of exogenous variables, and a vector $\boldsymbol{\eta}_i = \left(\eta_i^1, \eta_i^2, \eta_i^3\right)$ containing unobserved time invariant heterogeneity. Moreover, we denote by $g\left(p_{it}^*, r_{it}^*, w_{it}; \boldsymbol{\Sigma} | \boldsymbol{\eta}_i\right)$ the trivariate normal density, conditional on the random effects. Numerical approximation of the likelihood function proceeds in two steps. In the first step, the likelihood function is computed conditional on the unobserved individual characteristics. This first step

density is given by

$$f^C\left(\boldsymbol{y}_i|\boldsymbol{X}_i,\boldsymbol{\eta}_i;\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{\gamma},\boldsymbol{\Sigma},\alpha_{10}\right)$$

$$= \prod_{t=1}^{T_i}\int_{\mathbb{Q}_{it}}\int_{\mathbb{C}_{it}}\left\{(1-r_{it}^o)\left(1-\alpha_{10}\right)\int_{-\infty}^0 g\left(p_{it}^*,r_{it}^*,w_{it};\boldsymbol{\Sigma}|\boldsymbol{\eta}_i\right)\,\mathrm{d}r_{it}^*\right.$$

$$\left.+r_{it}^o\left[\int_0^\infty g(p_{it}^*,r_{it}^*,w_{it};\boldsymbol{\Sigma}|\boldsymbol{\eta}_i)\,\mathrm{d}r_{it}^*+\alpha_{10}\int_{-\infty}^0 g\left(p_{it}^*,r_{it}^*,w_{it};\boldsymbol{\Sigma}|\boldsymbol{\eta}_i\right)\,\mathrm{d}r_{it}^*\right]\right\}\,\mathrm{d}p_{it}^*\,\mathrm{d}w_{it}$$

The case where outmigration is perfectly observed follows by setting α_{10} equal to 0. The sets $\mathbb{W}_{it}$ and $\mathbb{P}_{it}$ define the domain of integration over the wage and work spaces and vary over time as individuals make different choices in each period according to the following table

Integration domains in period t		
	$\mathbb{Q}_{it}$	$\mathbb{C}_{it}$
Work	$-$	$[0,\infty)$
Not Work	$(-\infty,\infty)$	$(-\infty,0]$

Income is integrated out in waves where individuals do not work. The integration domain for the work propensity follows from the work decision rule. In the second step, the unconditional likelihood function is obtained by integrating out the random individual effects over $\mathbb{R}^3$

$$f\left(\boldsymbol{y}_i|\boldsymbol{X}_i;\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{\gamma},\boldsymbol{\Sigma},\boldsymbol{\Omega},\alpha_{10}\right)=\int_{\mathbb{R}^3}f^C\left(\boldsymbol{y}_i|\boldsymbol{X}_i,\boldsymbol{\eta}_i;\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{\gamma},\boldsymbol{\Sigma},\alpha_{10}\right)h\left(\boldsymbol{\eta}_i;\boldsymbol{\Omega}\right)\,\mathrm{d}\boldsymbol{\eta}_i$$

where H denotes the trivariate normal cumulative distribution function with mean vector $\boldsymbol{0}$ and covariance matrix $\boldsymbol{\Omega}$.

To solve the numerical integration problem, we approximate the integral by a simulated mean: a sequence of $r=1,2,...,R$ i.i.d. draws $\boldsymbol{\eta}_i^{(r)}=\left(\eta_i^{1(r)},\eta_i^{2(r)},\eta_i^{3(r)}\right)$ is taken from the multivariate normal distribution H at a given value of $\boldsymbol{\Omega}$.[9] For each draw, the conditional likelihood function f^C is evaluated. The partial MSL estimator consists of replacing f by the simulated mean

$$\frac{1}{N}\sum_{i=1}^N\log\left[\frac{1}{R}\sum_{r=1}^R f^C\left(\boldsymbol{y}_i|\boldsymbol{X}_i,\boldsymbol{\eta}_i^{(r)};\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{\gamma},\boldsymbol{\Sigma},\alpha_{10}\right)\right]$$

The resulting estimator is inconsistent for fixed R but will be consistent if R tends to infinity with the number of observations N. If $\sqrt{N}/R\to 0$ and with independent drawings across individuals, the method is asymptotically equivalent to maximum likelihood (Train, 2003). Finally, we model labor market earnings and work decisions of Germans using similar specifications of equations (7) and (8) and estimate the parameters using the simulation techniques described above.

4 Data

The data used in this paper is taken from the public use file of the GSOEP and covers the 1985-1999 period. Until 1990, the GSOEP consisted of two samples, A and B. Sample A consists of households with German heads living in former West-Germany. Sample B consists of an sample of immigrants living in West-Germany coming from countries which had signed a bilateral migration agreement with Germany in the 1950s and 1960s namely Greece, Italy, Spain, Turkey and Yugoslavia.[10] Data on speaking fluency, integration feelings of immigrants, intended length of stay and remittances directed to their family living outside Germany were given in consecutive waves from 1984 until 1987. Starting in 1987, this information was gathered every other year. In order to keep constant the time period between observations, we have chosen to keep the 8 waves of the panel where detailed information on immigrants was available, each spanned by one year, starting in 1985 and ending in 1999. Following the literature measuring the economic assimilation rate (e.g. Borjas, 1999a), we restrict our attention to males between 18 and 64 years of age during the 1985-1999 period. Excluded from the sample are individuals who died during the observation period and individuals who gave incomplete information on any single variable entering the empirical model in any of the 8 waves. This leaves us with a sample of 1987 native Germans and 732 immigrants starting in 1985.

The identification approach presented in section 2 relies on the information contained in panel attrition. It becomes instructive to contrast the attrition pattern of our immigrant sample with that of Germans whose attrition cannot obviously be attributed to outmigration. Table 1 contains information on the number of individuals observed along with the percentage of the original 1985 sample who remains in a given wave.[11] 41.9% of Germans and 26.7% of immigrants have been interviewed successfully in all the waves. The attrition rate in a given wave is defined as the percentage of individuals not observed in the given wave but observed in the preceding wave. Over our sample period, an average of 11.6% of the remaining Germans and 17.2% of immigrants drop out of the panel every two years. In the case of Germans, outmigration is de facto not an issue. Assuming that the difference in attrition rates is due to outmigration, a back of the envelope calculation implies that we would expect the outmigration rate in our sample of immigrants to be 6% every two years, or 3% per year, a number which would be in line with those reported in the literature (see Borjas and Bratsberg, 1996). Of course, this calculation relies on the assumption that attrition in the immigrant population which is not due to outmigration is of comparable magnitude to that of natives. We will come back to this in section 5 which reports indirect evidence suggesting that this should indeed hold in Germany.

The top panel of Figure 1 shows the average monthly gross income for working immi-

grants and Germans over the period covered. In 1985, the mean income of Germans was 3,357 DM per month compared to 2,690 DM per month for immigrants, giving an income ratio of 1.25 favoring Germans. The mean wage differential remaind relatively steady until 1991, after which, the mean income differential widened even more between the two groups to reach a ratio of 1.34 in 1999, with Germans receiving an average monthly wage of 5,848 DM while immigrants were receiving 4,348 DM per month. The diverging economic progress of Germans and immigrants after 1991 is also reflected in the work frequencies. The bottom panel of figure 1 shows the sample frequencies of individuals working in the month preceding the interview. We can see that until 1991, labor force participation was very similar for both Germans and immigrants. After 1991, we observe a steady decline in the work frequencies for both groups. During that period, the percentage of immigrants working remained steadily below that of Germans. The severe drop in work frequencies for both groups coincides with the general deterioration of the labor market in the regions of former West-Germany. Table 3 gives the unemployment rate per year by state. With the exception of Berlin, all provinces experienced their lowest unemployment rate of the 1985-1999 period in 1991. After 1991, the unemployment rate has progressively risen apart from a slight fall in 1999 for most provinces.

If outmigration does occur at a systematic time in the life-cycle, it is likely to affect the age and years since immigration composition of our sample of immigrants. Table 2 gives variable descriptions and summary statistics for Germans and immigrants for the 1985 and 1999 waves. We see that both Germans and immigrants are, on average, a little less than 40 years of age in 1985 while the average age of the cohort increases to 45 years of age in 1999 for both groups. As the average age of the sample increased only 8 years over this 14 year period, it is clear that both the relatively older Germans and immigrants left the panel. For a given mean age, Germans have acquired relatively more years of education, but relatively lower labor market experience than immigrants. The average number of months of labor market experience of immigrants increased by a little less than 32 months compared to 70 months for Germans, which is consistent with the fact that the proportion of working immigrants relative to Germans fell dramatically in the 1990's.

Most immigrants migrated to Germany early in their productive lives, a fact reflected by an average age at immigration of nearly 24 years, a figure consistent throughout the observation period which indicates that most migrants were old enough to autonomously decide to move to Germany. Both the evolution of years since immigration and immigration year are consistent with the hypothesis that outmigration occurs 20 years after migration (OECD, 2001). The average year of immigration of our cohort is 1969 in the

1985 wave, but raises to 1979 in the 1999 wave, indicating that the earlier cohorts are most susceptible to have dropped out of the panel. As the earlier cohorts contain the migrants with the higher number of years since migration in 1985, it is not surprising to see that average years since immigration increases relatively less than the 14 year time span, passing from 15.75 in 1985 to 19.63 in 1999, indicating again that earlier cohorts are those who left the panel. Reported feelings on integration in the German society and reported speaking fluency improved over time while health satisfaction seems to deteriorate, the latter likely capturing an aging effect. Finally, 73% of immigrants reported having a spouse living outside Germany in 1985 while as little as 1% still do so in 1999. This result can be interpreted in two ways. First, spouses may have eventually migrated to Germany during the time period. Second, it might be that immigrants whose spouse was living abroad were more likely to outmigrate.

5   Results and simulations

The regressors included in the earnings and work equations are education, labor market experience, labor market experience squared, self reported German speaking fluency, and the number of years since immigration to Germany. These are the standard variables that have appeared in this literature (Borjas, 1999a). The provincial unemployment rate in each wave is added in both equations to capture local labor market conditions. Finally, we include time fixed effects in each wave to capture remaining macroeconomic cyclical fluctuations. We use reported health satisfaction as the exclusion restriction in the work equation. Reported health satisfaction is a valid exclusion restriction if health problems occur mostly at a time in which an individual is more likely to have found a stable job whose continuation depends on the worker's choices. The outmigration equation includes as regressors whether or not the wife of immigrants lives in Germany, the immigrants self reported feelings of belonging to the Germans society and the unemployment rate. Immigrant's who arrive at a young age presumably have the highest incentive to invest in acquiring country specific human capital. This effect is captured by including age at arrival in Germany.

Immigrants investing in their native country presumably have different incentives to stay in the host country. The motives of these investments will affect their duration of stay. Immigrants may extend their stay in the host country in order to further increase their investments in their native country. On the other hand, if immigrants invest in starting up a business they wish to manage-as proposed by Dustmann and Kirchkamp (2002), high levels of investments will be associated with shorter migration durations.

To disentangle both hypothesis, we include in the outmigration equation the cumulative amount of money returned to the native country since 1984 as a proxy for investments. Reported health satisfaction and self-reported expected length of stay in Germany are also included, the later captures anticipatory behavior of migrants, which have been shown to affect the acquisition of country specific skills (e.g. Dustmann, 2002b). Time dummies are added to capture remaining macroeconomic fluctuations.

In order to separate the impact of selection on earnings into a work and outmigration effect, we first estimated an earnings equation with random effects. Our second specification is a bivariate model of labor market earnings and work. We finally estimated the complete model or earnings, work and outmigration. In the latter case, we experimented with an alternative specification of the outmigration equation which contained education, labor market experience and its square, speaking fluency, and years since immigration as regressors. A log-likelihood ratio test of the null hypothesis that these human capital variables have no joint effect on outmigration could not be rejected at conventional levels.[12]

5.1 Equation results

5.1.1 Covariance structure

We begin our analysis of the results with a discussion of the estimates characterizing the covariance structure of the unobserved components which are informative of the selection mechanisms. Table 4 presents estimates of the covariance structure. Focusing on the most general model which controls for both work and outmigration selection, we find a small but significant positive correlation between the unobserved individual heterogeneity of work and earnings $\left(\rho_{1,2}^{\eta}\right)$, indicating that individuals with higher probabilities of working are also more likely to have higher earnings. The correlation between individual time invariant heterogeneity of outmigration and earnings $\left(\rho_{1,3}^{\eta}\right)$ is -56% while that between outmigration and work $\left(\rho_{2,3}^{\eta}\right)$ is -49.8%, both significant at the 1% level. Both correlations suggest that individuals with a higher propensity to outmigrate are those with both a lower probability of finding work, and a lower labor market earnings, which points to negative outmigration selection. When comparing results with the bivariate model which does not correct for outmigration, we find that the estimated value of $\rho_{1,2}^{\eta}$ remains stable. Results for Germans are similar to that of the immigrant sample, with a small but positive and significant work selection effect $\left(\rho_{1,2}^{\eta}\right)$. Finally, transitory shocks between earnings and work, and shocks between work and outmigration, are all significantly negatively correlated, the former at -34.2% and the latter at 30.4% while we do not find significant correlation between the transitory shocks of the earnings and outmigration processes.

5.1.2 Outmigration

Table 5 presents the parameter estimates of the outmigration equation. We find that immigrants whose wife lives with them in Germany have a significantly lower probability of outmigrating, reflecting a preference for family unity. Immigrants satisfied with their health are significantly less likely to outmigrate, a finding consistent with the sociological findings reported in Stark (1998). Intended length of stay captures the expectations of immigrants and offers direct information on their remigration intentions. Not surprisingly, we find that migrants who expect to remain longer in Germany are also less likely to outmigrate. Deteriorations of the local labor market conditions, reflected in higher unemployment rates, have a positive and significant effect on the likelihood of outmigrating. The effect of cumulative savings returned to the home country is not significant, implying that immigrants who have saved and returned more money to their native country are not more likely to outmigrate. Dustmann and Kirchkamp (2002) find that Turkish return migrants have accumulated enough wealth in Germany to start up businesses in their home country upon their return. Because Turks are the biggest ethnic group in our sample, we would expect that increasing remittances increase the probability of outmigration if the money returned is intended to be eventually invested in a business. Our results suggest that migrants returned money to help relatives rather than for investment purposes. Finally, immigrant's feelings of integration in the German society are not correlated with outmigration.

In our data, the average attrition rate par wave (every two years) is 11.6% in the sample of Germans and 17.2% in the immigrant sample. If the proportion of immigrants dropping out of the panel and remaining in Germany is of the same magnitude to that of Germans, than the difference of 3% in annual attrition rates between natives and immigrants would be an estimate of the outmigration rate. We do not have direct information indicating that immigrants in Germany have the level of attrition which is not due to outmigration than natives. However, apart from outmigration and deaths, panel attrition occurs either because individuals decide to stop participating in the survey project, or individuals move within Germany and cannot be tracked by the survey institution. If Germans and immigrants have similar response rates or similar mobility, we would expect both groups to be equally well traceable. Clark and Drever (2001) have shown that immigrants in the GSOEP sample are not more likely to move within Germany than natives. Furthermore, Pischke and Velling (1997) have shown that immigrants in the western parts of Germany live in regions with a high concentration of ethnic minorities. Both results imply that, if anything, immigrants are easier to track than natives. Hence the proportion of immigrants dropping out and staying in Germany should be of similar magnitude to

that of Germans. This suggests that α_{10} should be no greater than 11.6%, which is what we find in the data.

We showed in section 2 that the conditional work probability and earnings of outmigrants are nonparametrically identified under the assumption that there exists a sample of immigrants with a probability of outmigrating close to 0. Estimates of our parametric model were used to compute the predicted probability that an immigrant leaves the host country in the following year. The 25th percentile of the outmigration probability distribution was 0.69% in 1985 and 0.81% in 1997, indicating that there is a considerable amount of immigrants with an outmigration probability close to 0.

5.1.3 Earnings equation

Table 6 presents parameter estimates of the earnings and work equations for the immigrant and German samples for all three models considered. We will first focus on the more general model which controls for both work and outmigration selectivity and subsequently compare the results with the more parsimonious models. Part of the earning disparity between Germans and immigrants can be explained by different returns to human capital investments. Returns to education of immigrants are roughly one third those of Germans, where an extra year of schooling raises earnings of Germans by 9.2% and those of immigrants by 2.9%. Part of the earnings differential between Germans and immigrants can also be related to returns to labor market experience. An extra year of labor market experience raises earnings of Germans by 0.6% compared to 0.32% for immigrants. The quadratic term of labor market experience indicates that the differential in earnings attributed to different returns to labor market experience diminishes as the level of experience increases. As expected, immigrants with better speaking fluency have higher earnings.[13] Finally, higher unemployment rates are associated with lower earnings in the immigrant population (at the 10% significance level) but do not affect labor market earnings of the native population. Given we include controls for time periods, identification of this effect relies on regional variations across provinces in Germany. The effect of unemployment on earnings thus reflects that at any point in time, earnings differ across provinces depending on the tightness of the local labor market.

When controlling for outmigration selectivity, we control for the fact that the sample of immigrants observed over time has above average conditional earnings relative to the population expectation. We then expect that the returns to some human capital factors will be lower in the overall immigrant population than in the population of permanent migrants. Changes in the returns to education when controlling or not for outmigration goes along those lines. We find that the returns to education of immigrants passes from

17

2.9% when outmigration is accounted for to 3.5% when we do not control for outmigration. This change is consistent with the hypothesis that permanent migrants have above average labor market earnings. The coefficient of years since migration progressively decreases as less selection is accounted for, passing from 0.079 to 0.073 when controlling for work selectivity to 0.069 without any selection controls, although these changes are not significant. Finally, the coefficient of the linear term of labor market experience increases while the coefficient of the quadratic term decreases when outmigration is not accounted for. Again, both these changes are not significant at conventional levels.

The robustness of parameter estimates to controls for work and outmigration selectivity has implications for estimates of the economic assimilation rate. The latter is generally defined as the differential in earnings between immigrants and natives with similar characteristics which results in one extra year in the host country

$$\left.\frac{\partial E\left(w_{it}|\boldsymbol{x}_{it}\right)}{\partial t}\right|_{immig} - \left.\frac{\partial E\left(w_{it}|\boldsymbol{x}_{it}\right)}{\partial t}\right|_{Germans}$$

For Germans, the passage of time is modelled as an increase of one year of labor market experience while for immigrants, the passage of time is modelled by increasing both the number of years of labor market experience and the number of years since immigration by one year. We computed the assimilation rate at the sample average of the labor market experience of immigrants in 1985. The estimate of the assimilation rate is found to be -5.00% per year when outmigration is not accounted for, and raises to -4.78% when outmigration is accounted for, a statistically insignificant increase. These results show that the conditional convergence of labor market earnings of immigrants to that of natives is not taking place in Germany. It is important to stress that the robustness of assimilation rates to outmigration selectivity is not inconsistent with our finding that outmigrants are negatively selected. The effect of outmigration on the estimated values of the assimilation rate will depend on both the earnings and work probability gap between outmigrants and immigrant stayers and on the size of the outmigration population. In the next section, we will present simulations which indicate that the earnings and work probability gaps between outmigrants and immigrants who stay are of sizeable magnitude, but the size of the outmigration population is small. Hence, assimilation rates should be relatively robust to outmigration selection.

5.1.4 Work equation

The results for the work equation are in line with those of the earnings equation, both in terms of the sign of the effects and on the robustness of the parameters to outmigration selection. For both immigrants and Germans, all parameters are statistically significant.

18

Education and labor market experience have positive effects on the probability of working. Higher unemployment rates have a negative effect on the work probability while immigrants and natives with better reported health satisfaction have a higher probability of working. Speaking fluency has a positive effect on work participation while the number of years since immigration has a negative and significant impact on work participation. Similar to earnings, we find that parameter estimates of the immigrant work propensity are quite robust to return migration selectivity. The only noticeable difference concerns the negative effect of the state level unemployment rate, whose coefficient passes from -0.054 to -0.076 when controlling for outmigration.

5.2   Simulations

Simulations are used for two purposes. First, they allow to check whether our model provides a good fit to the data. Secondly, they allow us to quantify the implications of outmigration selectivity on log monthly earnings and work propensities of immigrants.

Our simulations were done in the following way. For each individual appearing in the sample in 1985, we take 1000 draws from the joint distribution of the time invariant components $\left(\eta_i^1, \eta_i^2, \eta_i^3\right)$. Then, in each time period, we draw for each immigrant appearing in that period 1000 draws from the joint distribution of transitory stochastic components $\left(\varepsilon_{it}^1, \varepsilon_{it}^2, \varepsilon_{it}^3\right)$. The draws from the time invariant stochastic components are then matched to the draws of the transitory stochastic components and used to predict whether the immigrant will work and whether he will outmigrate in the next period. We compute the predicted log earnings for each immigrant predicted to work in a given wave. Simulations are then averaged over all draws and individuals. Simulations for the German sample follow a similar path. Simulation results are presented in Table 7. The fit for the German sample is very good, with both real and simulated paths closely following each other over the entire sample period. Simulated log earnings paths of immigrants are good up till 1991, after which, the model tends to over predict the monthly log earnings. Part of these discrepancies can be attributed to the progressively small immigrant sample sizes in the latter years, a fact reflected in the increasing dispersion of the simulated log earnings estimates over time (not shown here).

The empirical results of the previous section indicated that outmigrants were selected from the bottom of the earnings and work propensity distributions of the immigrant population. To gain some insights into the economic performance gap between immigrants who remained in Germany and those who left, we took the simulations which were used to compute results for immigrants in Table 7 and separated them into a group of predicted outmigrants and a group of predicted stayers. The top panel of Figure 6 reports simula-

tions of log earnings while the bottom panel reports the simulated work propensities. In 1985, the log earnings of outmigrants where 17.2% lower than those of the immigrant stayers. This gap widened to 20.1% in 1991 before dropping back to a gap of 17.7% in 1997. The gap in work propensities between immigrant stayers and outmigrants also confirms the strong negative outmigration selectivity. Outmigrants are predicted to have a work propensity 25% lower than that of immigrant stayers in 1985. In 1995, at the end of the economic downturn, this gap climbed to nearly 44%. Overall, these simulations clearly indicate that the economic performance of outmigrants in the wave preceding their departure was dramatically worse than that of stayers, both in terms of earnings and work probabilities.

6 Conclusions

This paper has presented a general framework to analyze the determinants of outmigration, to test the robustness of measures of economic assimilation, and to measure the parameters characterizing the economic performance of outmigrants using panel data subject to sample attrition. We estimated a three equation panel data model where outmigration decisions, work decisions, and expected labor market were jointly determined and allowed to depend on a rich set of observable characteristics, and on individual specific unobserved heterogeneity.

The model was estimated using the German and Immigrant samples of the Public use file of the GSOEP. The econometric results are broadly consistent with a pattern of negative outmigration selection. The model predicts an annual outmigration probability of 3% in our sample population. Simulations revealed that compared to immigrants who remain in Germany, outmigrants have between 17% and 20% lower expected labor market earnings, and between 25% and 44% lower probabilities of working over the 1985-1999 period. Interestingly, we do not find evidence that outmigration flows biases estimated assimilation rates.

Computation of the results above did not require that outmigration decisions were observable. The strategy used to identify the relevant parameters characterizing the economic performance of outmigrants relied on using sample attrition as a proxy variable for outmigration, and subsequently separating true outmigration movements from non-response which is unrelated to departures. Conditions for nonparametric identification of the work and outmigration probabilities, and the expected earnings of outmigrants have been presented. These conditions require that an economic model be specified in which some immigrants have an outmigration probability approaching zero. Importantly, as-

suming the econometric model is well specified, the validity of these conditions can be tested after estimation.

The economic model estimated in this paper is obviously just one example. Numerous directions of extension exist. The framework can be extended to any country where out-migration flows pose substantive concerns to policy makers, given survey data on these immigrants is available. The approach can also be extended in order to estimate economic structural static and dynamic models, allowing to recover preference parameters and to perform interesting policy simulations. Finally, the framework can be used more generally to study migration movements other than those of immigrants. Burda, Hardle, Muller, and Werwatz (1998)? for example study migration of native Germans from the East to West-Germany following reunification using data on intentions to migrate rather than actual migration movements, the later which were difficult to observe in their panel data set. The framework proposed here suggests that an extension to analyze such migrations is promising.

| | West-Germans | | | Immigrants | | |
	N	% 1985	Attrition rate %	N	% 1985	Attrition rate %
1985	1987	100	-	732	100	-
1987	1648	82.9	17.1	583	79.6	20.4
1989	1408	70.8	14.6	473	64.6	18.9
1991	1253	63.1	11.0	416	56.8	12.1
1993	1122	56.4	10.5	355	48.4	14.7
1995	1002	50.4	10.7	291	39.7	18.0
1997	919	46.3	8.28	242	33.1	16.8
1999	834	41.9	9.25	195	26.7	19.4
Mean 1985-1999			11.6			17.2

Table 1: Panel attrition for West-German and Immigrant samples 1985-1999. This excludes refreshment sample observations occurring in 1991 for Germans and in 1995 for immigrants.

	West-Germans		Immigrants		Variable Description
	1985	1999	1985	1999	
Age	38.33	45.85	39.78	44.53	
Experience	257.11	329.76	301.77	333.34	Number of months
Education	11.51	12.10	9.34	10.04	Number of years
Health satisfaction	7.26	6.81	7.21	6.77	0= unsatisfied,...,10 totally satisfied
Wife in Germany			0.73	0.01	1 if yes, 0 otherwise
Integration feeling			3.94	2.93	Do you feel German ?, 1=Totally, ...5=Not at all
German speaking fluency			2.65	2.30	1=excellent, 5=bad
Intended length of stay			2.18	0.59	1 = Within 1 year,2 = After a few years,3 = Never
Age at immigration			24.03	24.90	
Years since immigration			15.75	19.63	
Immigration year			1969	1979	
Number Obs.	1987	1264	732	393	

Table 2: Descriptive statistics for West-German and Immigrant samples, 1985 and 1999

23

State	1985	1987	1989	1991	1993	1995	1997	1999
Berlin	10.3	10.5	9.8	10.6	12.3	13.6	17.3	17.7
Schleswig-Holstein	10.8	10.3	9.6	7.3	8.3	9.1	11.2	10.6
Hamburg	12.0	13.6	11.7	8.7	8.6	10.7	13.0	11.7
Lower-Saxony	11.7	11.4	10.0	8.1	9.7	10.9	12.9	11.8
Bremen	14.5	15.6	14.6	10.7	12.4	14.0	16.8	16.8
North Rhine-Westphalia	10.7	11.0	10.0	7.9	9.6	10.6	12.2	11.2
Hess	7.0	6.7	6.1	5.1	7.0	8.4	10.1	9.4
Rhinel-Palatinate-Saarl.	9.4	9.1	7.8	6.1	8.3	9.2	11.0	9.7
Baden-Wuerttemberg	5.2	5.1	4.5	3.7	6.3	7.4	8.7	7.3
Bavaria	7.7	6.6	5.7	4.1	6.4	7.0	6.7	7.1

Table 3: Unemployment Rate per Wave and Land 1985-1999 Source: Statistisches Bundesamt Deutschland

	Immigrants			West-Germans
Work	No	Yes	Yes	
Outmigration	No	No	Yes	
$\rho^\varepsilon_{1,2}$		-0.416	-0.342	-0.153
		(0.103)	(0.119)	(0.069)
$\rho^\varepsilon_{1,3}$			-0.106	
			(0.107)	
$\rho^\varepsilon_{2,3}$			-0.304	
			(0.151)	
$\rho^\eta_{1,2}$		0.044	0.045	0.057
		(0.009)	(0.011)	(0.007)
$\rho^\eta_{1,3}$			-0.560	
			(0.143)	
$\rho^\eta_{2,3}$			-0.498	
			(0.173)	
$\sigma^2_{\eta^1}$	0.068	0.061	0.042	0.118
	(0.006)	(0.003)	(0.002)	(0.003)
$\sigma^2_{\eta^2}$		2.739	2.829	2.405
		(0.313)	(0.354)	(0.164)
$\sigma^2_{\eta^3}$			2.937	
			(0.373)	
σ^2_w	0.033	0.037	0.037	0.074
	(0.0005)	(0.0008)	(0.0008)	(0.0005)
Log-likelihood	-7137.81	-333.54	-701.76	-1372.04

Table 4: Covariance structure of the time variant and time invariant components. Asymptotic standard errors in parentheses.

Constant	-0.538
	(0.504)
Health satisfaction	-0.057
	(0.027)
Wife in Germany	-0.692
	(0.174)
Expected length of stay	-0.528
	(0.108)
Integration feeling	0.041
	(0.059)
Unemployment rate	0.069
	(0.029)
Age at immigration $/10^2$	1.618
	(0.755)
Cumulative remittances $/10^3$	-0.011
	(0.016)
α_{10}	0.107
	(0.013)

Table 5: Estimation results for outmigration. Asymptotic standard errors in parentheses. Wave dummies were included for 1987-1997 but are not reported in the table.

	Immigrants					West-Germans		
	Earnings			Work		Earnings		Work
Equation								
Work	No	Yes	Yes	Yes	Yes	No	Yes	Yes
Return migration	No	No	Yes	No	Yes	-	-	Yes
Constant	7.121	7.0198	7.063	-0.596	-0.532	5.971	6.030	-2.652
	(0.066)	(0.055)	(0.051)	(0.489)	(0.499)	(0.051)	(0.038)	(0.238)
Education / 10	0.257	0.346	0.288	1.007	1.060	1.002	0.915	1.489
	(0.045)	(0.029)	(0.029)	(0.326)	(0.327)	(0.031)	(0.016)	(0.130)
Experience / 10^3	3.112	3.150	3.239	21.934	22.002	5.999	6.087	31.613
	(0.201)	(0.174)	(0.175)	(1.484)	(1.612)	(0.147)	(0.157)	(0.966)
Experience squared / 10^4	-0.439	-0.431	-0.451	-3.439	-3.536	-0.786	-0.799	-5.509
	(0.029)	(0.028)	(0.028)	(0.217)	(0.239)	(0.024)	(0.027)	(0.158)
Years since immigration /10	0.069	0.073	0.079	-0.438	-0.416			
	(0.015)	(0.012)	(0.013)	(0.091)	(0.096)			
German speaking fluency	-0.017	-0.015	-0.020	-0.256	-0.232			
	(0.006)	(0.007)	(0.007)	(0.063)	(0.065)			
Unemployment rate	-0.005	-0.007	-0.005	-0.054	-0.076	-0.0041	-0.0004	-0.068
	(0.004)	(0.003)	(0.003)	(0.022)	(0.026)	(0.0029)	(0.002)	(0.015)
Health satisfaction				0.129	0.132			0.069
				(0.016)	(0.017)			(0.011)

Table 6: Estimation results for income and work equations. Asymptotic standard errors in parentheses. Wave dummies were included for 1987-1999 but are not reported in the table.

27

	Immigrants				Germans			
	Log earnings		% Working		Log earnings		%Working	
	R	S	R	S	R	S	R	S
1985	7.85	7.85	0.84	0.84	7.98	7.98	0.81	0.81
1987	7.93	7.94	0.85	0.86	8.10	8.09	0.83	0.83
1989	8.06	8.09	0.85	0.84	8.22	8.22	0.85	0.85
1991	8.16	8.20	0.82	0.82	8.31	8.31	0.86	0.86
1993	8.23	8.29	0.78	0.77	8.46	8.44	0.85	0.84
1995	8.29	8.38	0.73	0.71	8.50	8.49	0.85	0.84
1997	8.30	8.37	0.70	0.70	8.58	8.57	0.81	0.80

Table 7: Real (R) and simulated (S) log earnings of workers and work propensities for Germans and immigrants.
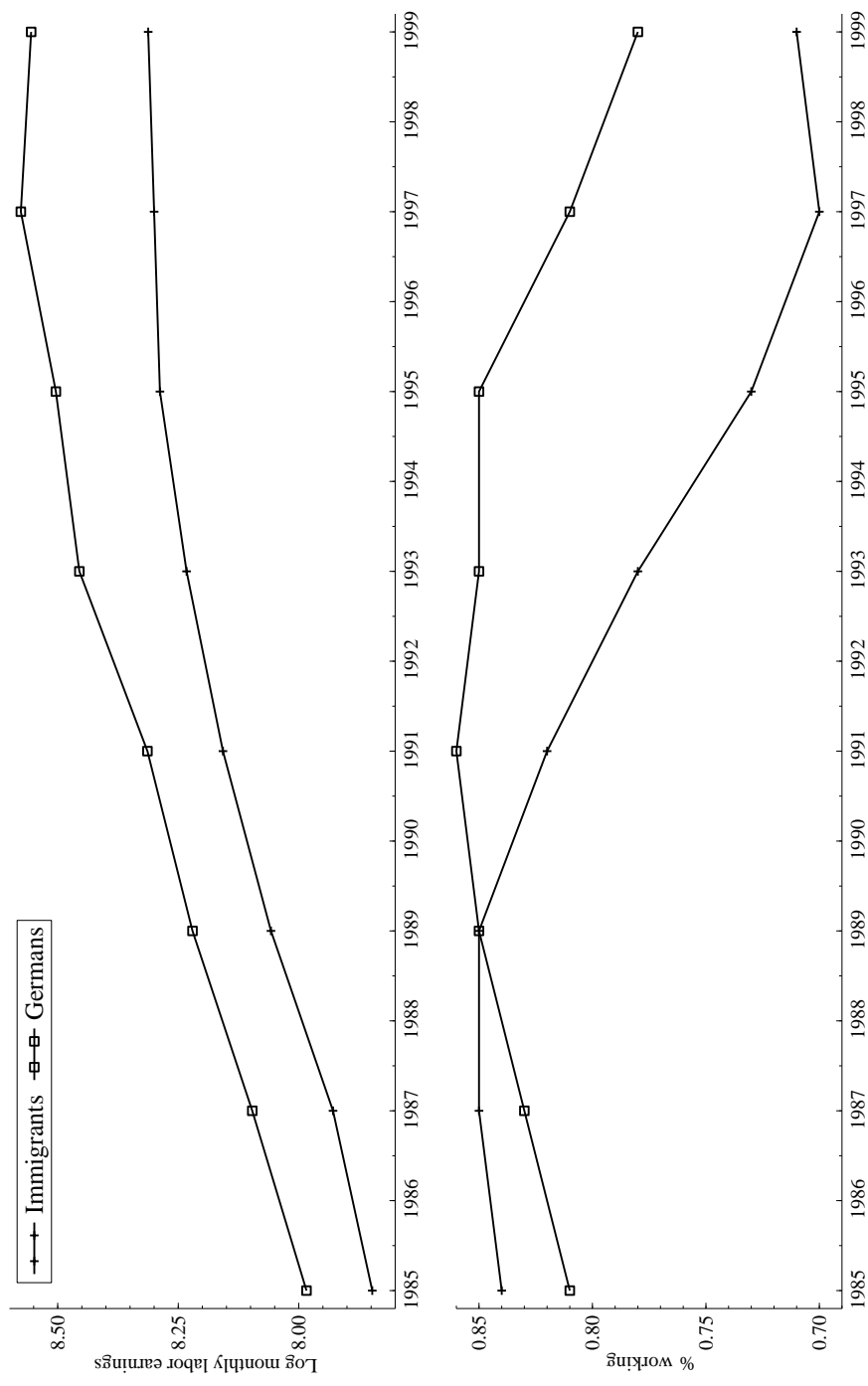
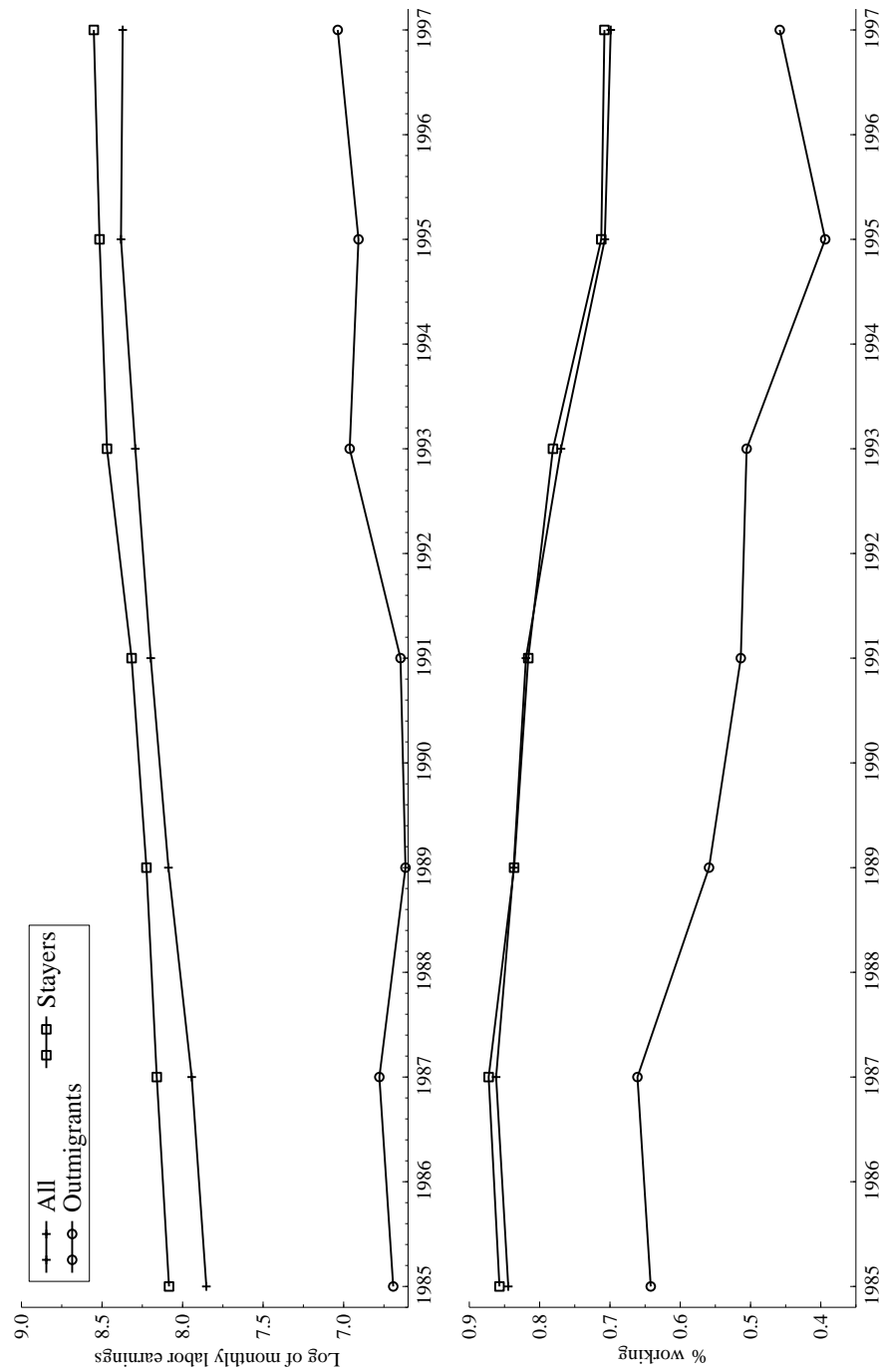Figure 1: Log monthly earnings and work propensities - Germans and immigrants

Figure 2: Top panel: simulated log earnings for immigrants over the 1985-1997 period. All, Stayers and Outmigrants refer to simulations averaged respectively over all immigrants, predicted stayers only and predicted outmigrants only. Simulations in each period are obtained by taking for each i 1000 draws from the distribution of $\varepsilon_{it}^1, \varepsilon_{it}^2, \varepsilon_{it}^3$ and $\eta_i^1, \eta_i^2, \eta_i^3$ and averaging over all draws the predicted earnings of those predicted to work. Bottom panel: Simulated proportion of immigrants working in the 1985-1997 period. Simulations are performed as in the top panel.

References

ARELLANO, M., AND B. HONOR (2001): *Panel Data Models: Some Recent Developments* vol. 5 of *Handbook of Econometrics*, chap. 53, pp. 3229–3297. Elsevier Science B.V.

BORJAS, G. J., AND B. BRATSBERG (1996): "Who Leaves ? The Outmigration of the Foreign-Born," *Review of Economics and Statistics*, 78, 165–176.

CARRINGTON, W. J., E. DETRAGIACHE, AND T. VISHWANATH (1996): "Migration with Endogenous Moving Costs," *American Economic Review*, 86(4), 909–930.

CLARK, W. A., AND A. I. DREVER (2001): "Do Immigrants Improve their Housing Quality When They Move ? Evidence from the German Socio-Economic Panel," *Vierteljahrshefte Zur Wirtschaftsforschung*, 70, 87–94.

DJAJIC, S., AND R. MILBOURNE (1988): "A General Equilibrium Model of Guest Worker Migration," *Journal of Internation Economics*, 25, 335–351.

DUSTMANN, C. (1993): "Earnings Adjustment of Temporary Migrants," *Journal of Population Economics*, 6, 153–168.

————— (2002): "Temporary Migration and Economic Assimilation," *Swedish Economic Policy Review*, 7, 213–244.

DUSTMANN, C., AND O. KIRCHKAMP (2002): "The Optimal Migration Duration and Activity Choice After Re-Migration," *Journal of Development Economics*, 67, 351–372.

EDIN, P.-A., R. J. LALONDE, AND O. ASLUND (2000): "Emigration of Immigrants and Measures of Immigrant Assimilation: Evidence from Sweden," *Swedish Economic Policy Review*, 7, 163–204.

HARRIS, J. R., AND M. P. TODARO (1970): "Migration, Unemployment, and Development: A Two-Sector Analysis," *American Economic Review*, 60, 126–142.

JASSO, G., AND M. R. ROSENZWEIG (1990): "Self-Selection and the Earnings of Immigrants: Comment," *American Economic Review*, 80(1), 298–304.

OECD (2001): "Trends in International Migration," Discussion paper, OECD.

PISCHKE, J.-S., AND J. VELLING (1997): "Employment Effects of Immigration to Germany: An Analysis Based on Local Labor Markets," *Review of Economics and Statistics*, 79, 594–604.

SCHULTZ, T. P. (1998): "Immigrant Quality and Assimilation - A Review of the US Literature," *Journal of Population Economics*, 11(2), 239–252.

STARK, O. (1995): "Return and Dynamics: The Path of Labor Migration When Workers Differ in their Skills and Information is Asymmetric," *Scandinavian Journal of Economics*, 97(1), 55–71.

——— (1998): *On the Microeconomics of Return Migration*chap. 3, pp. 32–41.

TRAIN, K. E. (2003): *Discrete Choice Methods with Simulation*. Cambridge University Press.