SOEPpapers
on Multidisciplinary
Panel Data Research

DIW Berlin
German Institute
for Economic Research

The German
Socio-Economic
Panel Study

30

Michael Fertig
Stefanie Schurer

Earnings Assimilation of Immigrants in Germany: The Importance of Heterogeneity and Attrition Bias

SOEPpapers on Multidisciplinary Panel Data Research
at DIW Berlin

This series presents research findings based either directly on data from the German Socio-Economic Panel Study (SOEP) or using SOEP data as part of an internationally comparable data set (e.g. CNEF, ECHP, LIS, LWS, CHER/PACO). SOEP is a truly multidisciplinary household panel study covering a wide range of social and behavioral sciences: economics, sociology, psychology, survey methodology, econometrics and applied statistics, educational science, political science, public health, behavioral genetics, demography, geography, and sport science.

The decision to publish a submission in SOEPpapers is made by a board of editors chosen by the DIW Berlin to represent the wide range of disciplines covered by SOEP. There is no external referee process and papers are either accepted or rejected without revision. Papers appear in this series as works in progress and may also appear elsewhere. They often represent preliminary studies and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be requested from the author directly.

Any opinions expressed in this series are those of the author(s) and not those of DIW Berlin. Research disseminated by DIW Berlin may include views on public policy issues, but the institute itself takes no institutional policy positions.

The SOEPpapers are available at
http://www.diw.de/soeppapers

ISSN: 1864-6689

German Socio-Economic Panel Study (SOEP)
DIW Berlin
Mohrenstrasse 58
10117 Berlin, Germany

Contact: Uta Rahmann | urahmann@diw.de

Earnings assimilation of immigrants in Germany: The importance of heterogeneity and attrition bias

Michael Fertig* and Stefanie Schurer†‡

August 17, 2007

Abstract

Heterogeneity in the ethnic composition of Germany's immigrant population renders general conclusions on the degree of economic integration difficult. Using a rich longitudinal data-set, this paper tests for differences in economic assimilation profiles of four entry cohorts of foreign-born immigrants and ethnic Germans. The importance of time-invariant individual unobserved heterogeneity and panel attrition in determining the speed of assimilation is analysed. We find evidence for heterogeneity in the assimilation profiles and for robust assimilation profiles for two entry cohorts only. Omitted variables, systematic sample attrition and the presence of second generation immigrants in the sample influence the speed of assimilation, but do not change the overall picture.

Keywords: Unobserved heterogeneity, panel attrition, sample selection, fixed effects, migration. **JEL Classification**: I12, C23.

*RWI Essen and ISG Cologne

†Ruhr Graduate School in Economics. Address for correspondence: RWI Essen, Hohenzollernstrasse 1-3, 45128 Essen, Germany. Email: schurer@rwi-essen.de, Tel: +49-201-8149-508

‡The authors thank Thomas Bauer, Jan Brenner, Annika Frohloff, Christoph M. Schmidt, Mathias Sinning, Marcus Tamm, and participants of the 10^{th} International Metropolis Conference, the 20^{th} Conference of the European Society of Population Economics, the 7^{th} GSOEP User Conference, the 61^{th} European Meeting of the Econometrics Society, and the 18^{th} Meeting of the European Association of Labour Economists for helpful comments.

1 Introduction

Since the end of the Second World War, Germany has faced a multi-cultural experience of immigration (Bauer et al., 2005). In the period up to the 1970s immigrants have been actively recruited from Southern Europe to match increased demand for low skilled labour in Germany's postwar economic boom. Since the end of the guest-worker recruitment era in 1973, the ethnic composition of immigration to Germany has changed substantially. Today, Germany has a sizeable community of ethnic Germans who originate predominantly from Eastern European countries and the former Soviet Union. This is a relatively young group of immigrants who arrived during the late 1980s and early 1990s and who received German citizenship upon arrival. Another large group that has arrived since the 1990s are the politically persecuted and refugees of war. This group in itself is very heterogeneous, including sending countries as diverse as Serbia and Montenegro, Turkey, Algeria, the Democratic Republic of Congo, Nigeria, Afghanistan, Iran and Iraq. This change in ethnic composition among entry cohorts implies a respectable degree of heterogeneity in educational backgrounds and German language capabilities. Expectation about assimilation behaviour should differ for each group.

Assimilation behaviour is traditionally tested in the framework of the assimilation hypothesis, an idea originally proposed by Chiswick (1978). It states that immigrants suffer an initial earnings disadvantage upon arrival vis-à-vis comparable natives. With years of residence, the initial earnings gap is expected to disappear. According to this idea, immigrants experience a steeper experience-earnings profile than natives, because they invest more in country-specific human capital accumulation due to lower opportunity cost (Dulep and Regets, 1999). The greater the initial disadvantage, the greater the incentive to acquire country-specific knowledge and therefore the faster the assimilation process. For the US, empirical tests of the assimilation hypothesis suggest a catch-up of foreigners with comparable natives after 10 to 15 years (Chiswick, 1978; Carliner, 1980). Proponents of this view argue that those immigrants that exceed the average earnings of comparable natives are positively self-selected in terms of unobservable characteristics

(Chiswick, 1978). The assimilation hypothesis is criticized by Borjas (1985, 1995) who argues that the positive and significant coefficients on years of residence typically estimated from cross-sectional data only capture a secular decline in the productivity of later immigrant cohorts. To support his argument, Borjas (1985) uses synthetic cohorts, that is following over time samples defined by year of immigration and age. The latter approach is however problematic in itself as it cannot capture selective migration, changes in the composition of samples over time or disentangle longitudinal changes from period effects (Chiswick et al., 2002).

The assimilation hypothesis in the framework of cross-sectional or pooled analysis has been the guiding analytical framework in assessing the degree of economic integration of immigrants in Germany. The majority of empirical work tests for a concave or a linear[1] earnings assimilation profile as a function of years of residence of guest-workers assuming the assimilation profile to be homogeneous across cohorts. Various studies yield, however, different results. Bauer et al. (2005), Licht and Steiner (1994), Schmidt (1997) and Pischke (1992) conclude that earnings of immigrants do not assimilate to those of comparable German natives over time, despite a large initial earnings differential upon arrival. Schmidt (1993), for blue-collar workers, and Constant and Massey (2005) find evidence for assimilation that takes place somewhere between17 to 23 years, respectively. Initial earnings differentials are explained by education (Schmidt, 1997; Constant and Massey, 2005) or country of origin differences (Schmidt, 1992), the latter being a proxy for differences in the quality of origin-country education or labour market experience gathered in the home country. Others explain the speed of assimilation over time by proxies for the actual investment in destination country-specific human capital. Dustmann (1993, 1994) suggests that intended length of stay and language proficiency are good predictors of economic assimilation.

In total, it is unclear whether long-run assimilation processes exist and whether these

[1]Schmidt (1992) tests for both linear and quadratic specifications. Schmidt (1994a) applies a linear specification for explaining assimilation of ethnic Germans. Schmidt (1997) uses indicator variables for various periods of lengths of stay in Germany. This study finds no assimilation of various cohorts of length of stay, however it does find that there are no initial earnings differences between within groups of qualification and status

differ across immigrant cohorts. Also, little is known about the labour market assimilation of ethnic Germans. Two exceptions are Bauer and Zimmermann (1997) and Schmidt (1994b) who find that ethnic Germans either do not have initial earnings disadvantages or that they eventually reach earnings parity with German natives. Both studies, however, use data collected before 1997. Last, we know only little about the economic position of immigrants from other countries of origin who arrived during the 1990s.

Using 21 waves of the German Socio-Economic panel, we assess the long-term economic position in terms of earnings of immigrants vis-à-vis German natives. Exploiting the longitudinal nature of our data allows to control for time-invariant, individual unobserved heterogeneity such as ability or latent health. For instance, Toussant-Comeau (2004) stresses the importance of unobserved heterogeneity in determining occupational upward mobility of Hispanics in the US and estimates the assimilation coefficient, i.e. the parameter picking up earnings dynamics, with a random effects specification. However, a random effects approach is limited to adjusting standard errors only rather than truly controlling for unobserved heterogeneity. In this case parameter estimates of upward mobility are more efficient, but they are still biased due to omitted variable bias. A fixed effects specification would be more appropriate.

The size of our data enables us to test whether the results are driven by including quasi-second generation immigrants in our sample. These are identified as immigrants who arrived in Germany at very young age, whereas we exclude a priori all foreigners actually born in Germany[2]. Age at immigration might be important since immigrants arriving at a very young age in the host country are more likely to acquire destination country-specific human capital such as language skills and knowledge about entry requirements into local labour markets.

In spite of a variety of advantages, the use of long panel data-sets entail some problems. The longer the sequence of waves the more likely it is that individuals systematically drop out of the sample. Panel attrition may bias estimation results if the probability of leaving

[2]Gang and Zimmermann (2000) identify second generation immigrants as those who were either born in Germany or arrived in Germany no older than 16 years of age.

the sample, either due to non-response or migration, is systematically linked to labour market outcomes. The empirical literature finds evidence for significant selectivity in exiting behavior, even though biases are rather small (Ayala et al., 2006; Behr, 2004; Behr et al., 2003, 2005; Becketti et al., 1988; Crouchley et al., 2002; Hausman and Wise, 1979; Lillard and Panis, 1998; Zabel, 1998; Ziliak and Kniesner, 1998). The majority of studies investigate attrition bias for the US Panel Study of Income Dynamics (PSID) or the European Community Household Panel (ECHP). For the GSOEP, work is limited to early waves of the data-set and does not differentiate between immigrant sub-groups, except for Schmidt (1994b, p. 121) who finds that 43 percent of individuals in his sample of guest-workers drop-out due to non-participation. Rendtel (1990, 1995) suggest that the impact of socio-economic variables on the probability of participation disappears after the second wave in 1985. Pannenberg (2000) and Spiess and Pannenberg (2003) demonstrate that there is substantial attrition from the GSOEP due to refusal or moving abroad. In the context of economic assimilation only Licht and Steiner (1994) test whether panel attrition in the GSOEP is systematically linked to labour market outcomes of foreigners[3]. On the other hand, biased assimilation coefficients could also be the result of selective return-migration, a special case of panel attrition. Typical determinants of return-migration are duration of stay, education or labour income, and location of family in the host country (Brecht, 1994; Constant and Massey, 2005; Velling, 1994; Schmidt, 1994a). Even though it is impossible to identify the actual return-migration in the GSOEP, we can model the decision to move abroad or the decision to stay in Germany. The underlying idea is that immigrants have a greater probability to move out of Germany than German natives and that their decision to move is highly linked with their duration of stay, income and educational background.

As a point of departure, we test the heterogeneity of economic assimilation profiles by augmenting a standard earnings equation with period of entry (cohorts) indicators. Cohorts are distinguished on the basis of changing immigration regimes, i.e. immigrants

[3]This study models both the labour market participation decision and the return-migration decision of foreigners.

who arrived between 1955-1968, 1969-1973, 1974-1987, and 1988-2002 and ethnic Germans arriving between 1988 and 2002. To control for unobserved heterogeneity, the same specification is re-estimated by taking differences from the mean. Technically, this method identifies the earnings growth rates of the cohorts over time, but no longer their initial earnings differences. Then, we model two possible sources of attrition bias. The two probabilities of participating in the interview and staying in Germany are corrected with a two-step Heckman sample selection model, modelling the two decision processes simultaneously. These estimates are used to calculate inverse Mills ratios separately for all foreigner cohorts and German natives. The challenge of this procedure is to identify appropriate exclusion restrictions for all groups. Last, we test the sensitivity of our results with respect to the sample definition by eliminating all individuals who could be defined as second generation immigrants.

We find evidence for heterogeneity in the assimilation profiles across cohorts for annual earnings. The assimilation hypothesis is confirmed for two cohorts only. Time-invariant unobserved heterogeneity and systematic drop-out of the sample influence the estimated speed of assimilation of these two groups, but the impact is still statistically significant. A check of robustness with respect to the sample definition leaves results mainly unchanged.

The paper is organized as follows. Section 2 explains the econometric framework with particular focus on the model to control for panel attrition and the choice of exclusion restrictions. Data issues are addressed in Section 3. Section 4 presents the empirical results and a checks for robustness, and in section 5 we summarize the findings.

2 Econometric framework

2.1 Labour market outcomes

To compare the labour market outcome of foreigners relative to German natives we augment a standard Mincer equation of log earnings with years of residence and its square. Let

Y_{it} represent real annual gross earnings[4] for individuals $i = 1, \ldots, N$ and $t = 1, \ldots, T_i$ (unbalanced panel) and take the natural logarithm of the column vector $Y_i = [Y_{i1}, \ldots, Y_{iT_i}]'$:

$$\begin{aligned} \ln Y_i &= \alpha + \sum_g D_g \cdot \beta_{g0} + \sum_g D_g \cdot YoR_i \beta_{g1} + \sum_g D_g \cdot {YoR_i}^2 \beta_{g2} + \\ &+ X_i'\theta + W_i'\gamma + H_i'\pi + I_i'\psi + u_i, \end{aligned} \tag{1}$$

where from now on we consider each enlisted variable as a column vector of dimension $T_i \times 1$ and each matrix of dimension $T_i \times k$, k being number of variables. In Eq. (1), 'years of residence' (YoR_i) measures the number of years a foreigner has resided in Germany after entry. The quadratic specification represents the assumption that log earnings are a concave function of years of residence (Chiswick, 1978)[5]. The matrix X_i includes a set of human capital dummy variables which take the value 1 if the individual holds a specific degree or vocational training, and 0 otherwise. We distinguish between five categories for schooling degrees, i.e. 'dropout', 'secondary schooling degree', 'intermediate schooling degree', 'technical schooling degree' and 'upper schooling degree', and four categories of professional training, i.e. 'no vocational training', 'vocation training', 'technocratic training', and 'university degree'. We define German natives to be the reference group captured by the constant α.

To allow the productivity to differ between immigrant cohorts (Borjas, 1985, 1995), the coefficients of the intercept β_{g0}, years of residence β_{g1} and its square β_{g2} vary across all four groups of first generation immigrants and ethnic Germans. The subscript g for group refers to Cohort 5568, Cohort 6973, Cohort 7487, Cohort 8802 and ethnic Germans. The dummy variable D_g equals one if the particular individual belongs to sub-group g, and zero otherwise. Years of residence and its square are interacted with each sub-group dummy.

[4] We use annual earnings since wages in Germany are relatively rigid. Employees have little influence on the wage determination process. Thus, wages do not necessarily reflect differences in labour market productivity. Lacking wage flexibility is particularly prevalent in the low skill sector, in which the majority of foreigners concentrate.

[5] We are aware of the critique by Murphy (1990) and Yuengert (1994) who show that a quadratic specification might not be the appropriate functional form. However, we chose the quadratic simplification since our main interest is to investigate the various sources of bias to the conventionally tested assimilation hypothesis by Chiswick (1978).

The matrix W_i includes a variety of individual-specific variables such as age, number of persons living in the household, marital status, and disability status. Workplace-specific variables such as the average hours worked per week and tenure at same firm are captured by the matrix H_i. We allow for hours worked to control for part-time employees. For immigrants, age at entry into Germany is captured by I_i for the sensitivity analysis only. All other determinants of earnings that cannot be observed are aggregated in the normally distributed zero mean error u_i. The main objects of interest in our analysis are the three parameter vectors β_{g0}, β_{g1}, and β_{g2}. Conditional on the specification, we impose various restrictions on the parameter vectors θ, γ, π, and ψ.

We do not include time fixed effects to capture business cycle variations. If included, we would have to make the assumption that foreigners and German natives are affected equally by business cycle shocks in order to identify the parameters (Borjas, 1994). This is the case because years of residence is a linear combination of the period effect and the year of immigration. Barth et al. (2004) have shown that equal period effect restrictions can produce biased estimates of assimilation and cohort effect, if the overall macroeconomic conditions have either a positive or negative trend. Since we use as dependent variable real earnings, that is adjusted to the price level, we can pick up inflation trends.

We also refrain from including self-assessed language proficiency as an explanatory variable as proposed by Dustmann (1994) and applied by Constant and Massey (2005). These subjective measures of language proficiency are prone to misclassification error and thus estimated coefficients may be severely biased (Dustmann and Van Soest, 2001). Moreover, language proficiency may be endogenous with respect to labour market earnings.

In a first step, we estimate Eq. (1) by pooled Ordinary Least Squares (POLS), imposing the restriction of a zero coefficient ($\psi = 0$) on the age at entry variable. In the pooled model we take advantage of the largest sample possible, which is particularly important given the small sample sizes of the immigrant sub-cohorts. In a second step, we re-estimate Eq. (1) with a linear fixed effects specification to address potential omitted variable biases due to time-invariant unobserved heterogeneity. Eq. (2) results from the

assumption $u_{it} = \alpha_i + \varepsilon_{it}$ and taking differences from the mean. Only parameters of time-varying variables can be identified.

$$\Delta \ln Y_i = \sum_g D_g \cdot \Delta YoR_i \beta_{g1} + \sum_g D_g \cdot \Delta YoR_i{}^2 \beta_{g2} + \Delta W_i' \gamma + \Delta H_i' \pi + \Delta \varepsilon_i. \qquad (2)$$

2.2 Panel attrition

The data used are unbalanced, the sequence of nonmissing observations varies across groups or even individuals. In this analysis we consider unit non-response only. On the one hand, individuals may refuse to participate any longer in the interview with no particular reason given. On the other hand individuals may drop out of the sample because they move abroad. If the underlying processes determining labour market outcomes correlates with those shaping the decision to participate or moving abroad OLS estimates are inconsistent (Heckman, 1979). For instance, assume that a disproportionately high share of low-skilled migrants compared to German natives leaves the panel prematurely due to language problems. If this group of low-skilled immigrants also exhibits a lower earnings potential than the individuals staying in the panel, OLS estimates of their speed of assimilation would be biased upward. Similar arguments hold for the decision to stay in Germany. For instance, if high-skilled migrants exhibit a higher probability to stay in Germany than low-skilled foreigners due to better labor market opportunities, OLS parameter estimates are biased upwards. If this systematic link between the two processes is constant over time, fixed effects estimation eliminates the bias. If not, even fixed effects estimation yields unreliable parameter estimates.

We address systematic panel attrition by assuming the existence of an unobserved variable that affects both the earnings equation and the attrition process. Under the assumption 'missingness on unobservables' (Fitzgerald et al., 1998) the bias can be alleviated with a Heckman sample selection model (Hausman and Wise, 1979; Verbeek and Nijman, 1992). Regarding the participation decision, we calculate the sample selection correction terms for German natives and ethnic Germans from a simple reduced form

probit model. For the different cohorts of first generation immigrants we calculate the selection correction from a bivariate probit model that links the error terms of the decision to participate in the interview and to stay in Germany. To identify the parameter estimates in the selection model we need good and valid exclusion restrictions.

For German natives and ethnic Germans let p_{ij}^* be the true, but unobserved net utility from participating in the interview:

$$p_{ij}^* = L.X_{ij1}\beta_{j1} + L.Z_{ij1}\gamma_{j1} + \epsilon_{ij1}, \tag{3}$$

where j represents these two groups, L is the lag operator, $L.X_{ij1}$ is a matrix of explanatory variables lagged by one time period, β_{j1} is vector of regression coefficients, and ϵ_{ij1} is an error term. The regressor matrix may coincide with all variables in Eq. (1). The vector $L.Z_{ij1}$ captures the exclusion restrictions lagged by one time period and the 1 in the subscript refers to the participation decision.

We observe the individual to participate in the interview, $p_{ij} = 1$,[6] if the true, underlying net utility from participating is greater than a threshold value, which we normalize to 0:

$$p_{ij} = \begin{cases} 1 & \text{if } p_{ij}^* > 0 \\ 0 & \text{if } p_{ij}^* \leq 0. \end{cases}$$

Assume $\epsilon_{ij1} \sim N(0,1)$ and let $(L.X_{ij1} + L.Z_{ij1})' = M_{j1}'$, $\theta_{j1} = (\beta_{j1}\ \gamma_{j1})'$, then the probability to participate can be expressed as:

$$Pr(p_{ij} = 1) = \Phi(M_{j1}'\theta_{j1}).$$

For German natives and ethnic Germans the inverse Mills ratio (IMR) can directly be estimated from a pooled probit model:

$$\hat{\lambda}_{j1}^1 = \frac{\phi(M_{j1}'\hat{\theta}_{j1})}{\Phi(M_{j1}'\hat{\theta}_{j1})}. \tag{4}$$

[6]The codification of this variable is based on the variable 'success of interview'. A value of1 in period t represents "successfully conducted interview in year t".

For the first generation sub-cohorts, we model the decision to participate and to stay in Germany jointly. The participation decision of foreign immigrants is analogous to (3):

$$p^*_{ic} = M'_{c1}\theta_{c1} + \epsilon_{ic1}, \tag{5}$$

c representing Cohort 5568, Cohort 6973, Cohort 7487, and Cohort 8802. For the decision to stay in Germany let s^*_{ci} be the true, but unobservable net utility from staying in Germany:

$$s^*_{ic} = M'_{c2}\theta_{c2} + \epsilon_{ic2}, \tag{6}$$

where $(L.X_{ic2} + L.Z_{ic2}) = M'_{c2}$ and $\theta_{c2} = (\beta_{c2}\ \gamma_{c2})'$. All variables are defined as above except for $L.Z_{ic2}$ being the vector of exclusion restrictions for this process and the subscript 2 refers to the decision to stay..

We observe a foreigner to stay in Germany $s_{ic} = 1$[7] if net utility from staying in Germany s^*_{ic} is greater than a threshold value, which we normalize to zero:

$$s_{ic} = \begin{cases} 1 & \text{if } s^*_{ic} > 0 \\ 0 & \text{if } s^*_{ic} \leq 0. \end{cases}$$

Assuming $\epsilon_{ic2} \sim N(0,1)$ the probability to stay in Germany can be expressed as

$$Pr(s_{ic} = 1) = \Phi(M_{c2}\theta_{c2}).$$

We further assume that the error terms of the two decisions are not independent from each other ($\text{cov}(\epsilon_{ic1}\epsilon_{ic2}) = \rho_c \neq 0$). The IMRs for the four different cohorts of first generation immigrants have to be calculated from a bivariate probit model in which we account for partial observability (Poirier, 1980; Vella, 1998). The error terms of (5) and (6) are

[7] The proxy for staying in Germany $s_{ic} = 1$ if the variable "success of interview" takes the value 5. This value represents moving out of Germany. For ethnic Germans we do not have to formalize this decision, because we obtain only four person-year observations for this group in our sample.

assumed to be distributed as:

$$\big(\epsilon_{ic1}, \epsilon_{ic2}\big) \sim \text{bivariate normal}\big(0, 0, 1, 1, \rho_c\big).$$

The log-likelihood is then:

$$\log L = \sum_{i=1}^{N} \log \Phi^{bp}(d_{ic1} M_{c1}\theta_{c1}, d_{ic2} M_{c2}\theta_{c2}, d_{ic1}d_{ic2}, \rho_c),$$

where $d_{icl} = 2y_{icl} - 1, l = 1, 2$ and Φ^{bp} is the bivariate normal cumulative distribution function. From this log-likelihood we obtain the bivariate probit Maximum Likelihood estimates $\hat{\theta}_{c1}$ and $\hat{\theta}_{c2}$ that are used to calculate IMRs according to Vella (1998, p. 256) for each cohort c:

$$\hat{\lambda}_{c1} = \sigma_{c1} \cdot \frac{\phi(M_{c1}\hat{\theta}_{c1})\Phi(M'_{c1}(\hat{\theta}_{c2} - \hat{\rho}_c \cdot \hat{\theta}_{c1}))}{\Phi^{bp}(M_{c1}\hat{\theta}_{c1}, M_{c2}\hat{\theta}_{c2}, \hat{\rho}_c)}, \tag{7}$$

and

$$\hat{\lambda}_{c2} = \sigma_{c2} \cdot \frac{\phi(M_{c2}\hat{\theta}_{c2})\Phi(M'_{c2}(\hat{\theta}_{c1} - \hat{\rho}_c \cdot \hat{\theta}_{c2}))}{\Phi^{bp}(M_{c1}\hat{\theta}_{c1}, M_{c2}\hat{\theta}_{c2}, \hat{\rho}_c)}, \tag{8}$$

The selectivity-corrected earnings equation is:

$$\begin{aligned} \ln Y_i &= \alpha + \sum_g D_g \cdot \beta_{g0} + \sum_g D_g \cdot YoR_i\beta_{g1} + \sum_g D_g \cdot {YoR_i}^2\beta_{g2} + \\ &+ X_i'\theta + W_i'\gamma + H_i'\pi + I_i'\psi + \sum_j \hat{\lambda}_j^{1\prime}\Gamma_j + \sum_c \hat{\lambda}_c^{1\prime}\Gamma_{c1} + \sum_c \hat{\lambda}_c^{2\prime}\Gamma_{c2} + u_i, \end{aligned} \tag{9}$$

where g refers to all groups, j to German natives and ethnic Germans, and c to first generation cohorts, 1 stands for the decision to participate and 2 for the decision to stay. The parameter vectors Γ_j, Γ_{c1} Γ_{c2} represent the influence of the inverse Mills ratios on earnings.

2.3 Exclusion Restrictions

A valid and good exclusion restriction, in our case Z, has to meet the following two conditions. It must not systematically correlate with the error term of the earnings equation ((i) $cov(L.Z_1, \epsilon_1) = cov(L.Z_2, \epsilon_2) = 0$) and it should significantly correlate with the participation or the decision to stay equation ((ii) $cov(L.Z_1, p) \neq 0$ and $cov(L.Z_2, s) \neq 0$). The first assumption requires that the exclusion restriction Z lagged by one time-period must not correlate with the unobservables that determine the current decision to participate in the interview or to stay in Germany. It implies that the exclusion restriction of last period must not correlate with current labour market outcomes. Whether this assumption holds has to be judged by economic reasoning. The second assumption requires that the exclusion restriction correlates with the decision to participate and to stay in Germany. It can be tested by imposing the Null-Hypothesis of $H_0 : \gamma_1 = \gamma_2 = 0$ in Eqs. (3), (5), and (6).

With respect to the participation decision, it is common to use 'change of interviewer during the first year since panel entry' (Behr, 2004; Rendtel, 1990; Spiess and Pannenberg, 2003; Willis and Hill, 2001). The idea behind this instrument is that interviewees are more likely to continue to participate if the interviewer remains the same over the year. Working through the questionnaire in collaboration with the interviewer is time intensive and to answer authentically requires trust towards the interviewer. If the interviewer changes, an interviewee must build up a new relationship, a requirement which may cause uneasiness. On the other hand, whether the interviewer changes does not influence the labour market performance of the interviewee, since this decision is solely taken by the data collection agency.

With respect to the decision to stay in Germany, it is more complex to find an appropriate exclusion restriction. The literature on return-migration identifies relative deprivation, capital constraints, higher purchasing power in destination country or country-of-origin, or higher rates of return to self-employment as possible explanations for returning home (see e.g. (Dustmann, 2003)). All of these factors are, nevertheless, intimately linked to

the labour market position of an immigrant. Constant and Massey (2005) suggest that any variable that represents strong ties or attachment with the country of origin is a good predictor for the probability of moving abroad. Information on where relevant family members live, whether the family has children in schooling age, or whether the immigrant came from a war-torn country may proxy these locational preferences. We choose indicators for 'number of children below the age 13', 'spouse or child(ren) away', and 'having left the country of origin due to war or seeking freedom'. The idea behind the exclusion restriction 'number of children below the age 13' is that families who have several children younger than 13 years of age are more likely to stay in Germany because they do not want them to change the familiar schooling environment.

Children who undergo primary and secondary education find themselves in a decisive period for developing intellectual and social skills. The more children of compulsory schooling age a family has, the more likely a family will decide in favour of staying. On the other hand, there is no empirical evidence that the actual number of children aged under 13 years exhibits an independent impact on the labour market position of the father[8]. Furthermore, a father whose child(ren) or spouse are living abroad will be more likely to return to the country where his family lives. Whether or not a part of the family stays abroad is more likely due to the particular immigration regime rather than due to the labour market outcome of the father.

Finally, whether an immigrant returns to his or her country-of-origin depends also on the motivation for migration. Immigrants who left their home countries to escape civil war or oppression of individual liberties are less likely to leave their host country as long as these conditions persist. On the other hand, whether or not such conditions are found in a specific country-of-origin is unlikely to be related to the labour market outcome of the particular immigrant.

[8]There are some arguments in favor of an existing link between the number of children aged younger than 13 and labour market earnings. The more children a family has, the more child benefits it receives. We use, however, gross annual earnings that exclude governmental transfers. On the other hand, the presence of children could motivate a family father to become more ambitious in his career. We found that the number of children in the time period before has no statistical significant association with contemporaneous earnings.

The former three instruments are used for the first three foreigner cohorts. In addition to these instruments, the instrument of war in country-of-origin is used for the latest immigrant cohort. Parameter estimates for both decision processes are statistically significant and are presented in Tables 7 and 8 in the Appendix.

3 Data

The analysis is based on 21 waves of the German Socio-Economic Panel (GSOEP) from 1984 to 2004. Our sample includes all male persons aged 18 to 60, who live in West Germany, who are not self-employed, and who are currently not in education or vocational training. This yields 86,510 person-year observations. German natives are identified as being born and raised in West Germany and holding a German citizenship. Ethnic Germans are identified as being born outside of Germany, holding a German citizenship, originating from Eastern Europe or Russia, and arriving in Germany after 1987. First generation immigrants are identified as being born outside Germany, entering Germany between 1955 and 2002, and holding a foreign nationality[9]. We further split this group of first generation immigrants into sub-cohorts which are identified along the various immigration regimes described in Bauer et al. (2005, p. 206-211).

We identify first generation immigrants who entered during the guest-worker recruitment period between 1955 and 1973. Since empirically a much larger number of immigrants entered after 1968 than between 1955 and 1967 (see Fig. 1 and Schmidt (1994a, p. 121)[10]) and since the last guest-worker agreement was signed in 1968, we distinguish between two groups of guest-workers. Those who entered between 1955 and 1968 are labelled Cohort 5568 and those who entered between 1969 and 1973 are labelled Cohort 6973. We opted for this sub-division of guest-workers to allow for a hypothetical systematic difference between the earlier and the later recruits in terms of risk attitude. The

[9] We disregard those immigrants who obtained the German nationality. More than1.5 percent of our total sample, or roughly 7 percent of the foreigner population in the sample, naturalized. Controlling for naturalization has no effect on the estimation results. These can be obtained on request.

[10] finds that more than 50 percent of his sample of guest-workers arrived in Germany between 1969 and 1973.

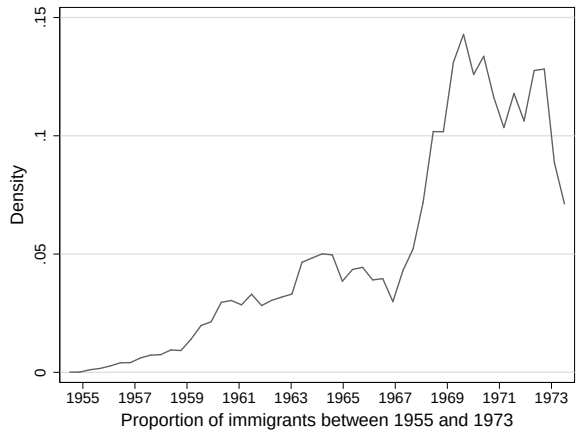


Figure 1: Inflow of immigrants between 1955 and 1973

earlier recruits may be interpreted as the pioneers who left their home country without social networks in Germany. The later recruits may have followed due to recommendations from fellow countrymen, who experienced the labour market opportunities, and due to the possibility of falling back on now existing social networks in Germany. Those immigrants who entered during family reunification between 1974 and 1987 are classified as Cohort 7487, and those who entered shortly before the fall of the Iron Curtain are classified as Cohort 8802. In total, we dispose of 9,977 native Germans, 297 ethnic Germans, and 2,152 first generation immigrants. For first generation immigrants we are left with 487, 749, 487, and 244 individuals for cohorts 5586, 6973, 7487, and 8802, respectively.

Table 1 presents the unconditional means of key socio-economic characteristics for all groups together with mean annual real gross labour earnings[11] (Tables 4 to 6 in the Appendix provide definitions for all variables). To account for the over-sampling of foreigners cross-sectional probability weights provided by the GSOEP are used. Longitudinal weights are not used since we explicitly model panel attrition. The vast majority of the first three foreigner cohorts stems from the classical guest-worker countries (71 to 82 percent), whereas this holds for less than 50 percent of the latest foreign immigrant cohort.

[11] This variable reflects the sum of all monthly salaries before tax deduction. It comprises bonus payments such as holiday bonus, and the so-called 13^{th} and 14^{th} monthly salary. We chose yearly income rather than monthly or hourly wages, since it captures times of unemployment or underemployment and represents the most important income concept in the German economy.

Hence, the nationality mix changed substantially in most recent years. Annual real gross earnings are the largest for native Germans, followed by Cohorts 5568, 6973, and 8802. Ethnic Germans and Cohort 7487 have the lowest earnings. For the former group it may be due to their members shortest duration of stay in Germany. For the latter group it is because its members are the youngest of all groups[12]. These distributional differences are exemplified in Fig. 2 for average earnings greater than $4,800$ Euro and smaller than $80,000$ Euro p.a.. The large fraction of individuals scattered around the 4800 Euro limit are those employed in the Mini-job sector[13].

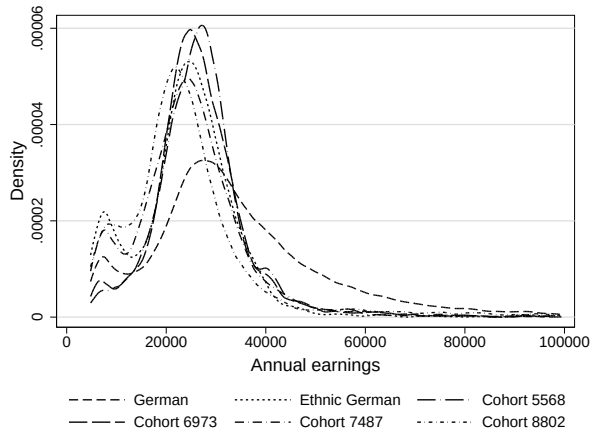


Figure 2: Probability distribution of annual earnings

The latest immigrant cohort is the group, which arrived on average at a much older age (26 years) than the three previous cohorts ($21, 22, 17$ years, respectively). Cohort 7487 was on average the youngest cohort to arrive in Germany. Nearly 42 percent of them arrived in Germany at age younger than 15 years. This is probably due to the fact that the main channel of immigration during that time was family reunification. In this

[12]The sample used does not include individuals with excessive real gross earnings for ethnic Germans. There were only six individuals whose real gross earnings exceeded $100,000$ Euro p.a.. Except for one individual, all hold a University degree and are older than 37 years of age. We excluded three cases from the German sample. Those were two individuals with less than 12.5 years of education and aged below 23 years, who earned more than $400,000$ Euro p.a. and one technocrat aged 30 years who reported to earn more than $500,000$ Euro p.a..

[13]Those are individuals working part-time or on the basis of tax-free income of currently 440 Euro per month. As the densities show, there is greater proportion of foreigners working in a Mini-job.

Table 1: Descriptive statistics

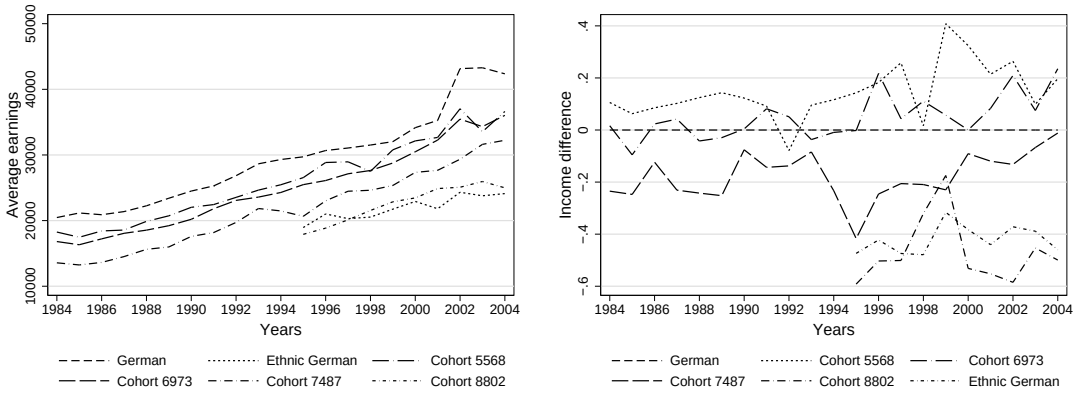
Group:	Age	% from main origin countries[1]	Age at immi-gration	% younger than 15 upon arrival	Highest schooling degree	No schooling degree	Annual real gross earnings[2]
German	38.9	−−	−−	−−	20.6	2.5	36,313
Ethnic Germans	36.4	91.9	27.8	15.9	4.4	5.1	24,426
Entry Cohort 55 to 68	48.5	71.8	22.0	18.6	2.2	31.2	27,509
Entry Cohort 69 to 73	44.5	81.8	21.1	26.0	2.3	24.8	26,948
Entry Cohort 74 to 87	34.0	70.8	17.0	41.9	3.0	14.3	22,701
Entry Cohort 88 to 02	34.1	47.8	25.9	7.5	4.8	9.5	24,836

Table 1 presents unconditional means of main variables of interests. [1] For first generation immigrants: Greece, Italy, Spain, Turkey and Yugoslavia. For ethnic Germans: Kazakhstan, Poland, Romania and Russia. [2] Annual gross earnings are averaged for annual incomes greater than 4,800 Euro.

group the majority of its members went at least partly through the German education system. Whether this makes any difference for their relative labour market performance remains an empirical question, which we address in our analysis. Education indicators suggest that this group is indeed different from earlier cohorts. A much smaller proportion of Cohort 7487 (around 14 percent) dropped out of school compared to the two oldest cohorts (32 and 24 percent). Ethnic Germans are relatively well educated, only 5 percent finished school without a degree and more than 4 percent hold the highest schooling degree. The youngest cohort of first generation immigrants is the group with the largest proportion of highly educated among foreign immigrants. However, we also observe a considerable share (nearly 10 percent) of its members without a schooling degree. Fig. 3(a) display the changes in income over time for all groups. Annual real gross earnings evolve heterogeneously over time for each sub-group relative to German natives. For instance, while earnings are growing for German natives, they are falling for nearly all foreigner groups (except Cohort 8802) between 1994 and 2004. Thus, we cannot make the assumption that time shocks equally affect foreigners and natives.

Fig. 3(b) displays the earnings differences between all foreigner sub-groups and the German benchmark-case (horizontal line)[14]. Income differences over time are small for

[14]This graph is the result of regressing the log of income or a indicator variable for unemployment on



(a) Growth

(b) Differences

Figure 3: Earnings over time

the first three cohorts. The youngest immigrant group, Cohort 8802, and ethnic Germans have the largest initial disadvantage, but their earnings increase the fastest over time. Income differences are statistically significant for all five sub-groups relative to native Germans (graph not presented here)[15]. They remain strictly negative for Cohort 7487, Cohort 8802 and ethnic Germans, and mainly positive for the oldest two Cohorts 5568 and 6973. The latter implies that we cannot observe a concave assimilation profile for these two groups in the raw data.

Figs. 4(a) to 4(f) show the evolution of the sample size for each group over time by professional training. Profession 1 means no professional training, profession 2 means the individual has acquired an apprenticeship, profession 3 means the individual has acquired a degree from a technical (three-year programme in administrative or health sciences) school, and profession 4 means the individual has obtained a university degree. Sample sizes change at different degrees for different professional groups and they change differently for the various sub-groups[16]. For all groups, sample sizes for individuals with a university degree diminish less strongly than for individuals with no professional training. For ethnic Germans, for instance, from 1999 onwards there are more individuals who hold

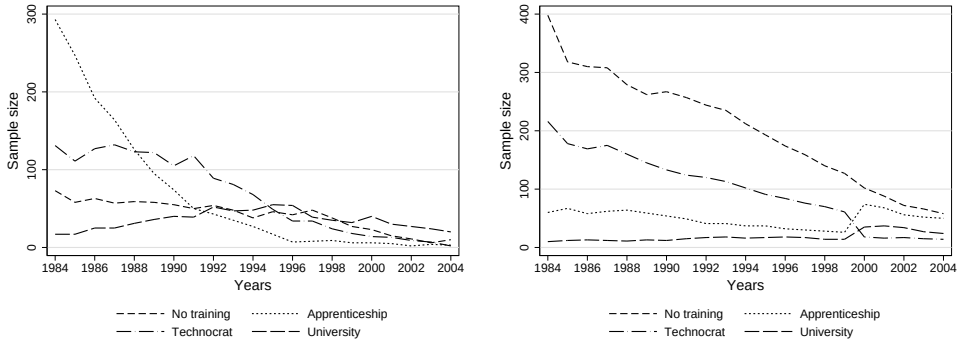
a set of dummy variables representing foreigner groups for each year. The parameter estimates of this raw method for each time period is used as data point.

[15]Graphs with confidence intervals are provided upon request.

[16]The large hikes in 1999 are the result of the refreshment samples added to the GSOEP in1999.

(a) German

(b) Ethnic German



(c) Cohort 5568

(d) Cohort 6973

(e) Cohort 7487

(f) Cohort 8802

Figure 4: Evolution of samples over time by vocational training

a university degree in the sample than those who have no professional training, whereas the opposite holds in 1994. Thus, there seems to be evidence for a systematic relationship between human capital endowment and the probability of staying in the sample.

4 Results

This section discusses the results obtained for the earnings equation. We report only marginal effects of interest to our economic hypothesis, i.e. the immigrant sub-group specific intercepts β_{g0}, which represent the initial earnings differential, and the parameter vectors of years of residence β_{g1} and its square β_{g2}. The latter indicate the assimilation profiles of each group. The results of the uni- and bivariate probit models, from which we calculate the inverse Mills ratios for the sample selection correction, are provided in Tables 7 and 8 in the Appendix. Full results are provided upon request[17].

4.1 Annual earnings

Estimates for Eq. (1) are presented in Table 2. In Model1 we regress the logarithm of real gross earnings on the sub-group indicators and the second order polynomial of years of residence. In Model 2 we extend this benchmark case by controlling for socio-demographic factors such as marital status, number of children, disability status, age and its square as well as the whole set of human capital indicators, i.e. type of secondary education and type of vocational training and workplace characteristics. Model3 estimates Eq. (9) correcting for panel attrition. These three models are estimated by pooled OLS (POLS). In Model 4 we estimate Eq. (2) to control for time-invariant unobserved heterogeneity. Finally, Model 5 combines Eqs. (2) and (9) assuming that the bias due to systematic drop-out of the sample is time-varying. In the latter two models we can only identify the coefficient of the assimilation profile for each sub-group separately and we cannot compare this profile to German natives. The last two models are therefore a robustness-check for

[17] The non-reported coefficients for the human capital indicators yield the expected signs and they are all statistically significant.

the estimates of the speed of assimilation.

Moving from Model 1 to Model 2 we replicate the important result from the literature that a large part of the initial earnings differences between foreigners and German natives can be explained by education and training differences. The only exception is Cohort 6973, but for all other groups the earnings gap decreases by at least 20 percentage points.

The most interesting case is the parameter estimates for ethnic Germans. Estimation results suggest a statistically significant concave earnings assimilation profile for ethnic Germans once controlling for education differences. This profile is remarkably robust across Models 2 to 5. Upon entry to Germany members of this group earn between 48 to 64 percent less than German natives[18]. However, their earnings grow substantially afterwards. In Model 2, for instance, four years of residence bring ethnic Germans a 10 precent increase in annual earnings. The quantitative extent of the initial earnings difference and catch-up to comparable natives critically depend on the chosen specification. The coefficient on years of residence decreases continuously from the full specification estimated by POLS to the Model with selectivity correction terms with individual fixed effects. Controlling for systematic panel attrition (moving from Model 2 to Model 3) decreases the initial earnings gap by nearly 16 percentage points. However, the speed of assimilation as captured by the estimate of the parameter on years of residence, decreases as well by 3.5 percentage point. The estimated coefficient for the quadratic term of years of residence remains unchanged. Further, controlling for omitted variables (moving from Model 2 to Model 4) has the same impact on parameter estimates. Only in Model 5 in which we control for both panel attrition and unobserved time-invariant heterogeneity the parameter estimates of the concave assimilation profile drop by more than 50 percent (Moving from Model 2 to 5)[19].

The latest foreign-born immigrant sub-group, Cohort 8802 also exhibits a statistically significant concave assimilation profile which is more robust across the various models

[18] In Euro terms that means an ethnic German earns about 16,000 Euros less than a German native annually in Model 1

[19] We tested whether the steep assimilation profiles of ethnic Germans are driven by individuals with extreme incomes. Excluding both the top 1 and 2 percent of income earners does not change the estimation results. These results are provided upon request.

Table 2: Results annual earnings

	Benchmark Specification POLS	Full Specification POLS	Full Specificcation Heckman correction	Full Specification Fixed Effects	Full Specification Heckman & Fix Eff
	Model 1	Model 2	Model 3	Model 4	Model 5
German	10.099***	6.277***	6.311***	6.145***	6.221***
	(.012)	(.097)	(.095)	(.042)	(.041)
IMR participation decision			.353***		.266***
			(.014)		(.006)
Ethnic German	-1.043***	-.886***	-.728***		
	(.291)	(.252)	(.240)		
Years of residence	.092	.152***	.117**	.113***	.063***
	(.057)	(.052)	(.051)	(.024)	(.024)
Years of residence^2	-.002	-.006**	-.005*	-.005***	-.003**
	(.003)	(.003)	(.003)	(.001)	(.001)
IMR participation decision			.467***		.392***
			(.077)		(.043)
Entry cohort 1955 **to** 1968	-.397	-.025	.364		
	(.473)	(.486)	(.507)		
Years of residence	.011	-.006	-.024	-.0009	.001
	(.035)	(.036)	(.038)	(.013)	(.013)
Years of residence^2	.0002	.0003	.0006	-.00002	-.00003
	(.0006)	(.0006)	(.0007)	(.0002)	(.0002)
IMR participation decision			.007		-.015
			(.056)		(.014)
IMR stay decision			.125		.194***
			(.170)		(.045)
Entry cohort 1969 **to** 1973	-.956***	-1.004***	-.116		
	(.214)	(.213)	(.172)		
Years of residence	.063***	.079***	.008	.028***	.016**
	(.019)	(.019)	(.016)	(.008)	(.008)
Years of residence^2	-.0007*	-.001**	.00004	-.0008***	-.0004***
	(.0004)	(.0004)	(.0004)	(.0002)	(.0002)
IMR participation decision			.234***		.058***
			(.024)		(.010)
IMR stay decision			-.440***		.132***
			(.100)		(.048)
Entry cohort 1974 **to** 1987	-.761**	-.050	-.040		
	(.340)	(.186)	(.191)		
Years of residence	.066*	.022	.020	-.002	-.014**
	(.039)	(.019)	(.019)	(.007)	(.007)
Years of residence^2	-.002	-.0007	-.0007	-.0005**	-.00008
	(.001)	(.0006)	(.0005)	(.0002)	(.0002)
IMR participation decision			.312***		.283***
			(.069)		(.020)
IMR stay decision			.051		.034
			(.073)		(.029)
Entry cohort 1988 **to** 2002	-1.436***	-.667**	-.616**		
	(.285)	(.274)	(.275)		
Years of residence	.202***	.118**	.102*	.116***	.062***
	(.062)	(.060)	(.059)	(.019)	(.020)
Years of residence^2	-.008**	-.005*	-.004	-.005***	-.003**
	(.003)	(.003)	(.003)	(.001)	(.001)
IMR participation decision			.087		.340***
			(.106)		(.040)
IMR stay decision			.536***		.169**
			(.194)		(.076)
Number of obs. $(N \cdot T)$	77879	67095	67095	67095	67095
R^2	.035	.437	.47	.282	.316
F statistic	58.651	173.708	167.796	794.709	689.412

Table 2 presents main estimation results on the assimilation profiles. IMR = inverse Mills ratio, semi-elasticities for the sub-group dummy variables are calculated as the difference (Δ_j) for each sub-group j vis-à-vis German natives as: $\Delta_j = \exp(\beta_{0,j}) - 1$ (Halvorsen and Palmquist, 1980). White robust standard errors are reported in parentheses. Significance levels are reported at1 %(***), 5 % (**) and 10 % (*)

than the one obtained for ethnic Germans. Controlling for the full set of education in Model 2 leaves this group with yet an initial earnings difference of nearly 50 percent to comparable German natives. Their earnings also grow with every additional year of residence, for instance, by 8 percent for four additional years of residence in Germany. This speed of assimilation is less strong when controlling for selective panel attrition in Model 3 and 5. Even though there are changes in the estimates for the assimilation profile when controlling for panel attrition (Moving from Model 2 to Model 3) or unobserved heterogeneity (Moving from Model 2 to Model 4), these are of minor economic importance. Again, only controlling for both potential biases simultaneously (Moving from Model 2 to Model 5) reduces the speed of assimilation substantially by nearly 50 percent.

By contrast, we cannot find a similarly robust picture of assimilation for Cohort6973, i.e. those individuals who immigrated during the last years of guest-worker recruitment. For them we observe a statistically significant concave assimilation profile in the full specification (Model 2), but once controlling for panel attrition (Model3), the concave earnings profile disappears. Taking into account time-invariant unobserved heterogeneity by individual fixed effects renders the coefficient for years of residence of this group statistically significant again. Controlling for selective panel attrition within the fixed effects framework does not change this result qualitatively, but it decreases the assimilation coefficient by almost 50 percent. Hence, in our sample the immigrant sub-group Cohort 6973 seems to be highly selected in terms of unobservables.

Both other cohorts, i.e. earlier guest-worker immigrants (Cohort5568) and immigrants from the family reunification era (Cohort 7487) do not exhibit any systematic assimilation profile. Especially in the case of Cohort 5568 we identify the assimilation profile mainly from individuals who have lived for many years in Germany. Thus, these are individuals who are close to the end of their assimilation process.

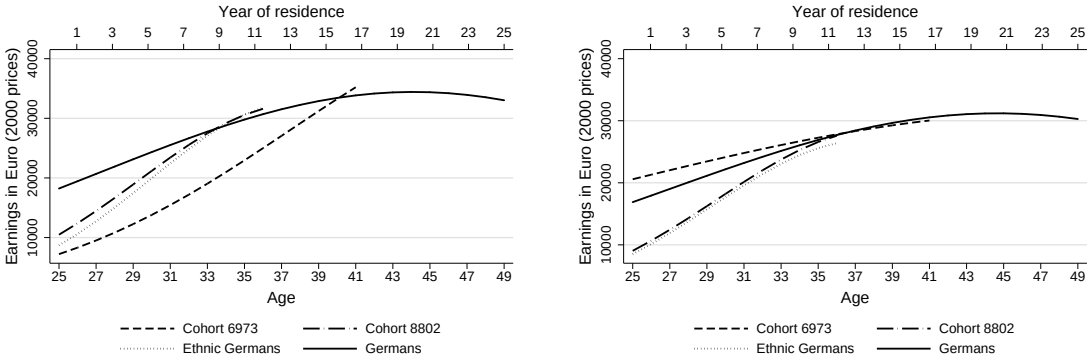
Last, the inverse Mills ratios (abrreviated as IMR in Table 2) for the participation decision are statistically significant and positive for all groups, except for Cohort5568 and Cohort 8802 in the POLS specification. This implies that unobservables that positively influence the earnings determination process also have a positive impact on the willingness

to participate in the interview. For all these groups income differences vis-à-vis German natives are exaggerated. On the other hand, the underlying links between the decision to stay and labour market outcomes are more ambiguous among the first generation cohorts. Unobservable factors that positively influence the labour market outcomes of Cohort8802 also positively influence the decision to stay. Even though Cohort5568 is also positively selected in terms of unobservables in the model with individual fixed effects (Model5), no statistical significant selection is found in Model3. In contrast, unobservables positively affecting the labour market performance of Cohort6973, negatively affect the decision to stay in Model 3. Our results suggest that the highly able members of Cohort6973 are more likely to leave Germany. By contrast, in the fixed effects specification of Model5 the highly able immigrants of Cohort 6973 are more likely to stay. One explanation for the alternating sign could be that the bias due to self-selection is time-varying, and that underlying factors such as personality traits affect the change of the bias over time.

Figs. 5(a) and 5(b) illustrate the differences in assimilation behaviour of ethnic Germans, Cohort 8802 and Cohort 6973 for Model 2 and 3, the latter being adjusted for selective sample drop-out. These profiles are constructed for an immigrant who arrived in Germany at age 25.[20].

Whereas both ethnic Germans and members of Cohort5568 catch-up with comparable German natives after less than 9 years, it takes members of Cohort 6973 nearly 16 years. Fig. 5(b) illustrates the small changes in assimilation behaviour after adjusting for panel attrition. The earnings of members of Cohort6973 are not statistically significant different from those of comparable Germans. In contrast, for the two remaining foreigner groups hardly any changes are visible except for the fact that earnings parity with German natives takes place more than two years later, if at all for Ethnic Germans.

[20]We constructed these profiles for an individual who obtained ten years of schooling, who holds an apprenticeship, who is married and has two children, who has been working for the past three years at the same firm, and who is fully employed (40 working hours per week)



(a) Full specification (Model 2) (b) Heckman corrected (Model 3)

Figure 5: Simulated assimilation profiles for earnings

4.2 Checks for robustness

Table 3 summarizes the main results of a checks for robustness to our preferred Model 2. Model S1 imposes the restriction of a homogeneous parameter vector of the assimilation vector across all foreigner groups. Here we test whether allowing the assimilation path to differ across entry cohorts matters. In Model S2 we exclude all Turkish immigrants from our sample. The restriction is justified by the hypothesis that it is mostly Turkish immigrants that face economic assimilation problems. In Models S3 to S5 we test whether it is mainly second generation immigrants which are driving the optimistic catch-up rates. Model S3 complements Model 2 with age at immigration as an additional regressor. Model S4 excludes all individuals who immigrated to Germany at an age younger than15. Model S5 excludes all immigrants whose potential labour market experience is smaller than their years of residence in Germany. This ensures that the sample includes only immigrants who did not undergo vocational training in Germany.

Not discriminating between entry cohorts for first generation immigrants and ethnic Germans (Model S1), we obtain an initial earnings difference much smaller than those obtained for ethnic Germans, Cohort 8802 and Cohort 6973, but much greater than for Cohort 5568 and Cohort 7487. A formal F-Test imposing parameter homogeneity for the initial earnings differences in Model 2 can be rejected at the five percent significance

level[21]. In Fig. 6(a) we can show that catch-up for foreigners as a homogeneous group would take place after less than 11 years.

Our results by and large do not change by excluding selective foreigner groups from the sample. Comparing Model 2 with Model S2, in which we exclude the Turkish population from our sample yields no significant change to your parameter estimates of interest. Fig. 6(b) shows that catch-up rates remain unchanged. Similar results are obtained when controlling for the second generation. Age at entry, even though statistically significant with a negative sign for all first generation entry cohorts, plays only a minor economic importance (Model S3). The only pronounced effect is observed for Cohort 6973, i.e. the second wave of guest-worker recruitment, whose members experience a 2.5 percent earnings penalty, all other things equal, for each additional year of age at immigration.

Ethnic Germans still exhibit a significantly concave earnings assimilation profile across all three models. Initial earnings differences vary within a range of no difference to around 11 percentage points when compared to Model 2. The quantitative dimension of the years of residence coefficient for ethnic Germans varies between 0.15 and 0.17 and is, thus, only slightly higher than in Model 2.

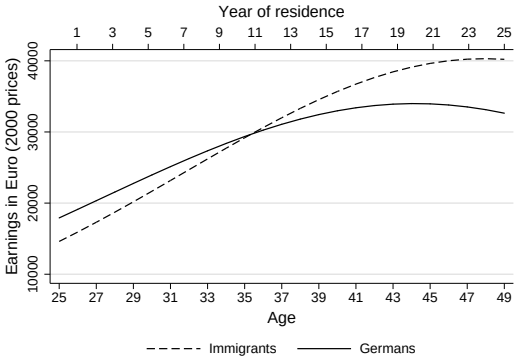
A slightly different picture emerges for Cohort 8802. This group still reveals a significantly concave earnings profile in Model S4, which is the sample excluding all immigrants who were younger than 15 years of age upon entry. Initial earnings differences grow by up to 15 percentage points in Model S4 and Model S5 vis-à-vis Model 2, but the speed of assimilation remains relatively constant across models. Fig. 6(c) exemplifies this change. The rapid catch-up rates for Cohort 8802 are mainly driven by the second generation. This group no longer reaches earnings parity with their native German peer group. For all other sub-groups the relevant estimates are insignificant and no clear pattern emerges.

[21] We tested jointly for equality of the cohort intercepts and the quadratic years of residence profiles after estimating the preferred Model 2: F(4, 10993) = 3.03, Prob > F = 0.0164 testing equality intercept estimates for foreigners; F(4, 10993) = 2.42, Prob > F = 0.0465 testing for equality of parameter vector of years of residence; and F(4, 10993) = 2.67, Prob > F = 0.0304 for the equality of the parameter vector of the square of years of residence.

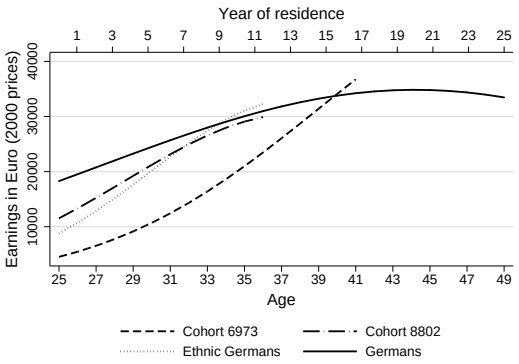
Table 3: Results annual earnings, sensitivity analysis

	Full specif	No cohort distinct	Exclude Turkish	Control for age at entry	No younger than 15 at entry	No immigs with training in GER
	Model 2	Model S1	Model S2	Model S3	Model S4	Model S5
Immigrant		-.246*** (.082)				
Years of residence		.025*** (.007)				
Years of residence^2		-.0004** (.0002)				
German	6.275*** (.097)	6.255*** (.108)	6.207*** (.111)	6.332*** (.098)	6.433*** (.100)	6.408*** (.100)
Ethnic German	-1.043*** (.252)		-1.040*** (.302)	-.877*** (.269)	-.930*** (.250)	-.941*** (.253)
Years of residence	.152*** (.052)		.163*** (.055)	.151*** (.054)	.166*** (.053)	.172*** (.053)
Years of residence^2	-.006** (.003)		-.006** (.003)	-.006** (.003)	-.007*** (.003)	-.008*** (.003)
Age at entry to Germany				-.002 (.004)		
Entry cohort 1955 **to** 1968	-.025 (.486)		.066 (.615)	.054 (.625)	.199 (.266)	.264 (.279)
Years of residence	-.006 (.036)		-.016 (.044)	-.002 (.036)	-.033* (.018)	-.038* (.020)
Years of residence^2	.0003 (.0006)		.0001 (.001)	.0002 (.0006)	.0009*** (.0003)	.0009*** (.0003)
Age at entry to Germany				-.008** (.004)		
Entry cohort 1969 **to** 1973	-1.004*** (.213)		-1.030*** (.336)	.126 (.256)	.043 (.189)	.026 (.196)
Years of residence	.079*** (.019)		.090*** (.028)	.034* (.018)	-.017 (.018)	-.017 (.019)
Years of residence^2	-.001** (.0004)		.001** (.0005)	-.0005 (.0004)	.0006 (.0004)	.0006 (.0004)
Age at entry to Germany				-.026*** (.003)		
Entry cohort 1974 **to** 1987	-.050 (.186)		-.225 (.295)	.140 (.151)	-.115 (.184)	-.075 (.198)
Years of residence	.022 (.019)		.036 (.033)	.016 (.017)	.022 (.020)	.004 (.021)
Years of residence^2	-.0007 (.0006)		-.001 (.001)	-.0007 (.0005)	-.0007 (.0006)	.00004 (.0007)
Age at entry to Germany				-.008* (.005)		
Entry cohort 1988 **to** 2002	-.667** (.274)		-.668* (.393)	-.517 (.394)	-.825*** (.288)	-.807*** (.288)
Years of residence	.118** (.060)		.114 (.088)	.124** (.060)	.143** (.063)	.123** (.062)
Years of residence^2	-.005* (.003)		-.006 (.005)	-.005* (.003)	-.006** (.003)	-.005 (.003)
Age at entry to Germany				-.010 (.007)		
R^2	.437	0.374	0.3936	.441	.430	.431
F statistic	173.708	244.61	162.93	160.04	154.17	152.80

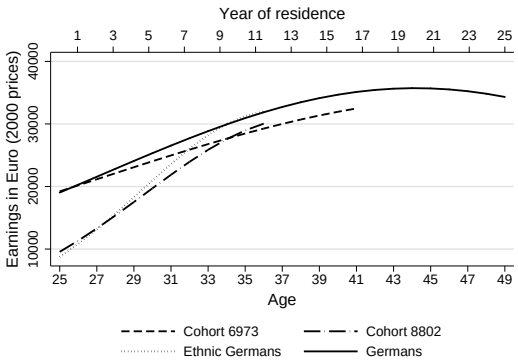
Table 3 reports the results on the robustness checks applied to the preferred model (earnings equation controlling for all socio-demographic variables including human capital indicators). Model S1 imposes homogeneous parameter vectors on assimilation profiles, Model S2 excludes all Turkish immigrants from the sample, Model S3 includes "age at immigration" as additional regressor, Model S4 restricts the sample to only those foreigners who arrived in Germany at age 15 years or older, and Model S5 excludes all foreigners who went, at least partly, through the German education system. White robust standard errors are reported in parentheses. Significance levels are reported at1 %(***), 5 % (**) and 10 % (*)



(a) Parameter homogeneity (Model S1)



(b) Exclude Turks (Model S2)



(c) Exclude second generation (Model S5)

Figure 6: Simulated assimilation profiles, robustness checks

5 Conclusions

This paper investigates labour market outcomes in terms of annual gross real earnings of immigrants relative to comparable German natives. Rich and high quality data of the German Socio-Economic Panel are utilized to test for heterogeneous assimilation profiles across four entry cohorts of first generation immigrants and ethnic Germans. We test whether the estimates of the assimilation profiles are driven by potential differences in unobserved productivity or human capital investments that determine the shape of the assimilation path between groups. In addition, we investigate the role of panel attrition in shaping the estimation results. Systematic exit from the sample is corrected in a two-step sample selection correction that models jointly the decision to participate in the interview

and the decision to stay in Germany (Heckman, 1979).

We find heterogeneity in the assimilation profiles across the five identified groups. With respect to annual earnings, ethnic Germans and the youngest group of foreign immigrants, both who arrived between 1988 and 2002, exhibit similar statistically significant concave assimilation profiles. Both groups suffer an initial disadvantage vis-à-vis comparable German natives of nearly 50 percent, whereas their earnings grow at a decreasing rate over time, making up approximately ten percent for each additional four years of stay in Germany. These initial earnings differences are greater than those found in Schmidt (1992) for guest-workers who predicts an initial disadvantage of earnings of twelve percent and an average increase of 0.7 percent for each additional year of residence. This indicates that more recent immigrants and ethnic Germans perform worse upon arrival, but assimilate reasonably well to the German labour market if they find employment. The estimated catch-up rates of both groups of less than nine years are also in-line with findings for US data (Chiswick, 1978; Carliner, 1980). For the foreigner group which arrived between 1969 and 1973, we find a flatter assimilation profile. Catch-up for this group occurs after 16 years, an estimate which resembles the results of Schmidt (1993) and Constant and Massey (2005). We cannot confirm the assimilation hypothesis for immigrants who arrived between 1955 and 1968 and between 1974 and 1987. Our estimates show neither statistically significant initial differences nor earnings growth rates in comparison to German natives. This might be due to the fact that we do not observe earnings differences upon entry from 1955 onwards for these cohorts, but rather upon sample entry since 1984. Thus, assimilation profiles are mainly identified with data points collected late along the assimilation path for Cohort 5568.

Our results further suggest that omitted variable and attrition bias play a minor quantitative role in these outcomes. Only for Cohort 6973 we find substantial differences, indicating that this group is highly self-selected. Unobserved time-invariant heterogeneity slightly biases the estimated assimilation profiles upwards. Nevertheless, controlling for this bias with fixed effects estimation still yields statistically significant coefficients on earnings growth. Regarding selective panel attrition, we find that for most foreigner sub-

groups and German natives alike unobservable factors that affect earnings positively, also impinge positively upon the decision to participate in the interview. Hence, individuals who perform relatively well in the labour market are also more likely to stay in the sample. This result on foreigners is in line with results provided by Rendtel (1995) for the general population and the general literature on return-migration for Germany (Brecht, 1994; Constant and Massey, 2005; Velling, 1994; Schmidt, 1994a). Controlling for attrition bias leaves earnings differences of ethnic Germans and Cohort 8802 vis-à-vis German natives less pronounced and assimilation profiles flatten out, even though the estimates are still significant. Only for Cohort 6973 systematic drop out from the sample seems to drive assimilation profiles. Taken together, our results correspond with the empirical literature on panel attrition (e.g. Behr (2004) or Behr et al. (2003)) suggesting that labour market related attrition is present, but does not necessarily alter conclusions.

Imposing a homogeneous parameter vector on the assimilation profile is rejected by conventional F-tests. Looking at the assimilation path, we find nevertheless that a homogeneous group of foreigners suffers an initial earnings difference of approximately twenty percent, but earnings parity with comparable German natives would be reached after eleven years. Excluding the quasi-second generation immigrants from our sample changes results only slightly for Cohort 8802. This entry cohort no longer reaches earnings parity with German natives. This suggests that for individuals from this group who arrived in Germany before the age of 14 or who obtained some vocational training in Germany perform relatively well economically. Last, we cannot find that Turkish immigrants perform significantly different from their immigrant peers, as excluding them from all foreigner sub-samples does not change estimation results.

Overall, we conclude that both observable and unobservable heterogeneity plays a significant role in assessing labour market outcomes of individuals with different cultural backgrounds. A separate analysis of the second generation could be of interest to complete the picture of heterogeneous assimilation behaviour.

References

Ayala, L., C. Navarro, and M. Sastre (2006). Cross-country income mobility comparisons under panel attrition: The relevance of weighting scheme. *ECINEQ Working Paper Series 2006-47*.

Barth, E., B. Bratsberg, and O. Raaum (2004). Identifying earnings assimilation of immigrants under changing macroeconomic conditions. *Scandinavian Journal of Economics 106*.

Bauer, T., B. Dietz, K. Zimmermann, and E. Zwintz (2005). German migration: Development, assimilation, and labour market effects. In K. Zimmermann (Ed.), *European Migration. What do we know?*, pp. 197–261. Oxford: Oxford University Press.

Bauer, T. and K. F. Zimmermann (1997). Unemployment and wages of ethnic Germans. *Quarterly Review of Economics and Finance 37*, 361–377.

Becketti, S., W. Gould, L. Lillard, and F. Welch (1988). The Panel Study of Income Dynamics after fourteen years: An evaluation. *Journal of Labour Economics* (6), 472–492.

Behr, A., E. Bellgardt, and U. Rendtel (2005). Extent and determinants of panel attrition in the European Community Household Panel. *European Sociological Review 21*(5).

Behr, A., E. Bellgart, and U. Rendtel (2003). The estimation of male earnings under panel attrition. A cross-country comparison based on the European Community Household Panel. *CHINTEX Working Paper 11*.

Behr, A. (2004). Comparing estimation strategies in the presence of panel attrition: Empirical results based on the ECHP. *mimeo*.

Borjas, G. J. (1985). Assimilation, changes in cohort quality, and the earnings of immigrants. *Journal of Labour Economics 3*(4), 463–489.

Borjas, G. J. (1994). The economics of immigration. *Journal of Economic Literature 23*, 1667–1717.

Borjas, G. J. (1995). Assimilation and changes in cohort quality revisited: What happened to immigrant earnings in the 1980s? *Journal of Labour Economics 13*(2), 201–245.

Brecht, B. (1994). *Remigration von Gastarbeitern - Statistische Modellierung von Rückkehrprozessen*. Dissertation: Universität Konstanz.

Carliner, G. (1980). Wages, earnings and hours of first, second, and third generation American males. *Economic Inquiry 18*, 87–102.

Chiswick, B. R., Y. L. Lee, and P. W. Miller (2002). Longitudinal analysis of immigrant occupational mobility: A test of the immigration assimilation hypothesis. *IZA Discussion Paper Series DP No. 452*

Chiswick, B. R. (1978). The effect of Americanization on the earnings of foreign-born men. *Journal of Political Economy 86*(5), 897–921.

Constant, A. and D. S. Massey (2005). Labour market segmentation and the earnings of German guestworkers. *Population Research and Policy Review 24*(5), 489–512.

Crouchley, R., S. Bradley, and R. Oskrochi (2002). Evaluating the impact of missing data in social reserach: Simulations and applications using the BHPS and the NCDS. *Lancaster University 8 January 2002. Mimeo.*

Dulep, H. O. and M. C. Regets (1999). Immigrants and human capital investment. *American Economic Review. Papers and Proceedings 89*(2), 186–191.

Dustmann, C. and A. Van Soest (2001). Language fluency and earnings equation with misclassified language indicators. *Review of Economics and Statistics 83*, 663–674.

Dustmann, C. (1993). Earnings adjustment of temporary migrants. *Journal of Population Economics 6*, 153–168.

Dustmann, C. (1994). Speaking fluency, writing fluency and earnings of migrants. *Journal of Population Economics 7*, 135–156.

Dustmann, C. (2003). Return migration, wage differentials, and the optimal migration duration. *European Economic Review 47*, 353–369.

Fitzgerald, J., P. Gottschalk, and R. Moffit (1998). An analysis of sample attrition in panel data - The Michigan Panel Study of Income Dynamics. *Journal of Human Resources 33*, 251–299.

Gang, I. N. and K. F. Zimmermann (2000). Is child like parent? Educational attainment and ethnic origin. *Journal of Human Resources 3*, 550–569.

Halvorsen, R. and R. Palmquist (1980). The interpretation of dummy variables in semi-logarithmic equations. *American Economic Review 70*(3), 474–475.

Hausman, J. and D. Wise (1979). Attrition bias in experimental and panel data: The Gary Income Maintenance Experiment. *Econometrica 47*, 455–474.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica 47*, 153–161.

Licht, G. and V. Steiner (1994). Assimilation, labour market experience, and earnings profiles of temporary and permanent immigrant workers in Germany. *International Review of Applied Economics 8*(2), 130–156.

Lillard, L. A. and C. W. A. Panis (1998). Panel attrition from the Panel Study of Income Dynamics: Household income, marital status, and mortality. *Journal of Human Resources 33*(2), 437–457.

Murphy, K. A. (1990). Empirical age-earnings profiles. *Journal of Labour Economics 8*(2), 202–229.

Pannenberg, M. (2000). Documentation of sample sizes and panel attrition in the German Socio-Economic Panel (GSOEP) (1984-1999). *DIW Discussion Paper 232*.

Pischke, J.-S. (1992). Assimilation and the earnings of guestworkers in Germany. *ZEW Discussion Paper Series 92-17*.

Poirier, D. J. (1980). Partial observability in bivaraiate probit models. *Journal of Econometrics 12*, 210–217.

Rendtel, U. (1990). Teilnahmeentscheidung in Panelstudien: Zwischen Beeinflussung, Vertrauen und sozialer Selektion. Über die Entwicklung der Antwortbereitschaft im Sozio-ökonomischen Panel. *Kölner Zeitschrift für Soziologie und Sozialpsychologie 42* 280–299.

Rendtel, U. (1995). *Panelausfälle und Panelrepräsentativität*. Frankfurt/New York: Campus Verlag.

Schmidt, C. M. (1992). Country-of-origin differences in the earnings of German immigrants. *Universität München Discussion Paper 92-29*.

Schmidt, C. M. (1993). The earnings dynamics of immigrant labour. *Centre for Economic Policy Research Discussion Paper No. 763*

Schmidt, C. M. (1994a). Country of origin, family structure and return migration. *Vierteljahrshefte zur Wirtschaftsforschung 63*, 119–125.

Schmidt, C. M. (1994b). The economic performance of germany's east european immigrants. *CEPR Discussion Paper 94-09*.

Schmidt, C. M. (1997). Immigrant performance in Germany: Labour earnings of ethnic German migrants and foreign guestworkers. *Quarterly Review of Economics and Finance 37*, 379–397.

Spiess, M. and M. Pannenberg (2003). Documentation of sample sizes and panel attrition in the German Socio-Economic Panel (GSOEP). *Deutsches Institut für Wirtschaftsforschung Research Notes 28*.

Toussant-Comeau, M. (2004). The occupational assimilation of hispanics in the US: Evidence from panel data. *Federal Reserve Bank of Chicago WP 2004-15*.

Vella, F. (1998). Estimating models with sample selection bias: A survey. *Journal of Human Resources 33*, 127–169.

Velling, J. (1994). The determinants of family reunification among german-guest-workers. *Vierteljahrshefte zur Wirschaftsforschung Proceedings of the 1993 International Conference of GSOEP Users*, 126–132.

Verbeek, M. and T. Nijman (1992). Testing for selectivity bias in panel data models. *International Economic Review 33*(3), 681–703.

Willis, R. and D. Hill (2001). Reducing panel attrition: A search for effective policy instruments. *Journal of Human Resources 36*(3), 416–438.

Yuengert, A. M. (1994). Immigrant earnings, relative to what? The importance of earnings function specification and comparison points. *Journal of Applied Econometrics 9*, 71–90.

Zabel, J. (1998). An analysis of attrition in the Panel Study of Income Dynamics and the Survey of Income and Program Participation with an application to a model of labor market behavior. *Journal of Human Resources 33*(2), 479–506.

Ziliak, J. P. and T. J. Kniesner (1998). The importance of sample attrition in life cycle labor supply estimation. *Journal of Human Resources 33*(2), 507–30.

Table 4: Description of variables

Variable Name	**Description**
Assimilation Variables	
Ethnic German	Intercept Dummy which takes the value 1 if the individual belongs to the group of ethnic Germans and 0 otherwise
Cohort 5568	Intercept dummy which takes the value 1 if the individual belongs to the group of First Generation Immigrant, Cohort 5568 and 0 otherwise
Cohort 6973	Intercept dummy which takes the value 1 if the individual belongs to the group of first generation immigrant, Cohort 6973 and 0 otherwise
Cohort 7487	Intercept dummy which takes the value 1 if the individual belongs to the group of first generation immigrant, Cohort 7487 and 0 otherwise
Cohort 8802	Intercept dummy which takes the value 1 if the individual belongs to the group of first generation immigrant, Cohort 8802 and 0 otherwise
Yrsres	Years of residence calculated as the difference between the year in time period t and the year of immigration
Yrsres Eth Ger	Interaction term of rears of residence and dummy ethnic German
Yrsres Cohort 5568	Interaction term of rears of residence and dummy Cohort 5568
Yrsres Cohort 6973	Interaction term of rears of residence and dummy Cohort 6973
Yrsres Cohort 7487	Interaction term of rears of residence and dummy Cohort 7487
Yrsres Cohort 8802	Interaction term of rears of residence and dummy Cohort 8802
Yrsres^2	The Square of Years of Residence calculated as the square of the difference between the year of current time period t and the year of immigration
Yrsres^2 Eth Ger	Interaction term of Years of Residence Squared and Dummy Ethnic German
Yrsres^2 Cohort 5568	Interaction term of Years of Residence Squared and Dummy Cohort 5568
Yrsres^2 Cohort 6973	Interaction term of Years of Residence Squared and Dummy Cohort 6973
Yrsres^2 Cohort 7487	Interaction term of Years of Residence Squared and Dummy Cohort 7487
Yrsres^2 Cohort 8802	Interaction term of Years of Residence Squared and Dummy Cohort 8802
Set of Education Dummies	
Dropout	Dummy which takes the value 1 if the individual dropped out of high school and 0 otherwise
Secondary	Dummy which takes the value 1 if the individual passed compulsory schooling of nine years of high school (Hauptschule) and 0 otherwise
Intermediary	Dummy which takes the value 1 if the individual received ten years of high school (Realschule) and 0 otherwise
Upper	Dummy which takes the value 1 if the individual received thirteen years of high school (Gymnasium) and 0 otherwise

Table 5: Description of Variables, continued

Variable Name	**Description**
Set of Professional Education Dummies	
No training	Dummy which takes the value 1 if the individual did not acquire any professional training, 0 otherwise
Training	Dummy which takes the value 1 if the individual finished an apprenticeship or vocational school of any profession (Lehre) , 0 otherwise
Technocrat	Dummy which takes the value 1 if the individual obtained a degree of a health care, technical or civil service training (Fachhochschule) ,0 otherwise
Uniedu	Dummy which takes the value 1 if the individual received a degree from a technical college or a university , 0 otherwise
Othschool	Dummy which takes the value 1 if the individual received any other education which is left unspecified , 0 otherwise
Socio-Demographic Charact.	
Age	Continuous variable that measures the current age of an individual
Age^2	Continuous variable that measures the square of age
Mar1	Dummy which takes the value 1 if the individual is married , 0 otherwise
Mar2	Dummy which takes the value 1 if the individual is married but lives separated from the partner , 0 otherwise
Mar3	Dummy which takes the value 1 if the individual is single , 0 otherwise
Mar4	Dummy which takes the value 1 if the individual is divorced , 0 otherwise
Mar5	Dummy which takes the value 1 if the individual is widowed , 0 otherwise
Disable	Dummy which takes the value 1 if the individual is registered as being disabled of any degree , 0 otherwise
Pershh	Number of persons living in the household
Working Life Charact.	
Workhrs	Continuous variable that measures the average weekly hours spent at work
Ten	Continuous variable that measures the number of years an individual spent at the firm currently working

Table 6: Description of Variables, continued

Variable Name	**Description**
IV Participation Decision	
Change	Change of Interviewer: This variable takes the value1 if the interviewer changed for the individual after the first interview
IV Non-Return Migration Decision	
Child13	Number of children in household which are younger than 13 years of age
Child Away	Dummy variable that takes the value 1 if the individual's children are living in home country
Spouse Away	Dummy variable that takes the value 1 if the individual's spouse lives abroad
War/Freedom	Dummy variable that takes the value1 if the individual's motivation to migrate to Germany was either to escape war or to search for political freedom

Table 7: Results of reduced form probit for decision to stay

	Cohort 55 to 68	Cohort 69 to 73	Cohort 74 to 87	Cohort 8802
	(1)	(2)	(3)	(4)
No. child. < 13 (IV)	.246***			
	(.063)			
Child away (IV)		.548***		
		(.180)		
Spouse away (IV)			-1.029**	
			(.440)	
War or freedom (IV)				.450*
				(.271)
Age	.359***	.088**	-.061	-.083
	(.047)	(.035)	(.043)	(.071)
Age^2	-.004***	-.001***	.001*	.001
	(.0005)	(.0004)	(.0006)	(.0009)
Workhrs per week	.006	.003	-.009	-.014*
	(.004)	(.005)	(.007)	(.008)
Tenure in firm (Yrs.)	.019***	.015**	.023***	-.063**
	(.005)	(.006)	(.009)	(.030)
Intermediate (10 Yrs.)	.842**	-.053	4.909***	4.934***
	(.338)	(.238)	(.166)	(.178)
Technical (10 to 12 Yrs.)		4.799***	-1.227***	
		(.221)	(.297)	
Upper (13 Yrs.)	.031	4.667***	.101	-1.517***
	(.388)	(.213)	(.261)	(.377)
Other	-.792***	-.868***	-.975***	-.565***
	(.148)	(.199)	(.124)	(.213)
Dropout	-1.046***	-.729***	-1.120***	-1.161***
	(.167)	(.203)	(.161)	(.373)
University degree	-.263	-.398	.132	-.201
	(.271)	(.257)	(.215)	(.280)
No prof. training	-.546***	-1.020***	.521***	-.155
	(.142)	(.211)	(.150)	(.281)
Technocrat raining	-.474***	-.977***	.420***	-.595**
	(.154)	(.216)	(.138)	(.267)
Separated	1.072***	-.420	-.453	6.321***
	(.287)	(.370)	(.336)	(.255)
Single	.430***	-.138	.631***	-.324
	(.157)	(.206)	(.133)	(.264)
Divorced	.959***	-.045	1.198***	6.671***
	(.225)	(.223)	(.231)	(.403)
Person in HH (No.)	.036	.119***	.211***	-.035
	(.039)	(.037)	(.036)	(.061)
Const.	-6.134***	1.137	2.210***	4.174***
	(1.060)	(.871)	(.837)	(1.304)
Obs. (N*T)	2869	5304	3007	855

Table 7 reports coefficients of a binary probit model that regresses the observation of not moving out of Germany on a set of regressors and instrumental variables (IV). Instruments used are the number of children below the age of 13 (No. child. < 13), Child away, Spouse away, and war or freedom. (N*T) = Number of person-year observations. Significance levels are reported at 1 %(***), 5 % (**) and 10 % (*).

Table 8: Results of reduced form probit for decision to participate

	GER	ETH GER	C 5568	C 6973	C 7487	C 8802
	(1)	(2)	(3)	(4)	(5)	(6)
Change of interviewer (IV)	-.412***	-.554***	-.550***	-.487***	-.838***	-.541***
	(.020)	(.215)	(.117)	(.087)	(.108)	(.184)
Age	-.002	-.057	-.023	.234***	-.033	.068
	(.006)	(.043)	(.054)	(.045)	(.034)	(.054)
Age^2	.0001*	.0006	.0007	-.002***	.0006	-.0006
	(.00007)	(.0005)	(.0006)	(.0005)	(.0005)	(.0007)
Workhrs per week	-.003***	-.005	-.012***	.023***	.008	-.001
	(.0009)	(.007)	(.005)	(.004)	(.005)	(.006)
Tenure in firm (Yrs.)	.003**	.024	.015***	.009*	.011	.010
	(.001)	(.016)	(.005)	(.005)	(.009)	(.021)
Intermediate (10 Yrs.)	.009	.230	-3.123***	1.296***	.114	-.026
	(.021)	(.165)	(.301)	(.309)	(.156)	(.349)
Technical (10 to 12 Yrs.)	.074**	-.380				
	(.037)	(.346)				
Upper (13 Yrs.)	.160***	-.278	-.691*	.664***	.747***	-.139
	(.032)	(.232)	(.353)	(.254)	(.205)	(.385)
Other	.017	.391***	-.169	.583***	-.083	-.234
	(.093)	(.134)	(.189)	(.166)	(.104)	(.241)
Dropout	.122*	.639	.384**	.534***	-.355***	-.010
	(.069)	(.501)	(.184)	(.166)	(.125)	(.324)
University degree	-.164***	.356**	-.802**	.038	-.556***	-.319
	(.032)	(.164)	(.322)	(.194)	(.176)	(.246)
No prof. training	-.099***	-.018	-.769***	.339**	-.049	-.315
	(.030)	(.145)	(.178)	(.154)	(.100)	(.220)
Technocrat Training	-.069***	-.019	-.557***	.328**	.064	.006
	(.021)	(.131)	(.192)	(.155)	(.147)	(.236)
Separated	-.117	-.930*	.781**	-.104	-.008	-1.413***
	(.073)	(.490)	(.356)	(.244)	(.325)	(.413)
Single	-.091***	-.047	1.409***	2.103***	-.050	-.653***
	(.025)	(.176)	(.190)	(.243)	(.115)	(.223)
Divorced	-.144***	-1.260***	.273	1.109***	-.184	-.982***
	(.037)	(.310)	(.182)	(.220)	(.177)	(.284)
Person in HH (No.)	.040***	.026	.136***	.267***	.112***	.192***
	(.007)	(.025)	(.030)	(.030)	(.021)	(.051)
Const.	-.051	1.362	-.331	-9.680***	-.298	-1.227
	(.126)	(.835)	(1.200)	(1.192)	(.625)	(1.004)
Obs. (N*T)	44864	1162	2869	5265	2942	855

Table 8 reports coefficients of a binary probit model that regresses the observation of participating in the interview on a set of regressors and instrumental variables (IV) lagged by one time period. Instrument used is the change of interviewer after the first year participating. GER = German, ETH GER = ethnic German, C5568 = Cohort 5568, C 6973 = Cohort 6973, and so on, (N*T) = Number of person-year observations. Significance levels are reported at1 %(***), 5 % (**) and 10 % (*).