THE IMPACT OF SKILLS ON REMIGRATION FLOWS

Peter Bönisch, Philipp Gaffert and Joachim Wilde

Working Paper 82
March 2010

INSTITUTE FOR EMPIRICAL ECONOMIC RESEARCH
University of Osnabrück
Rolandstraße 8
49069 Osnabrück
Germany

The Impact of Skills on Remigration Flows

Peter Bönisch

Martin-Luther-University Halle-Wittenberg
peter.boenisch@wiwi.uni-halle.de

Philipp Gaffert

Martin-Luther-University Halle-Wittenberg
philipp.gaffert@web.de

Joachim Wilde

University of Osnabrück
joachim.wilde@uni-osnabrueck.de

Abstract

More than ten years after the seminal paper by Borjas and Bratsberg (1996) modeling the impact of skills on remigration the empirical evidence on that theory is still mixed. Our paper is to shed light on that issue. Using the GSOEP we test two hypotheses derived from Borjas and Bratsberg (1996) while allowing for endogeneity of host country specific capital. Our results give strong support for their theory. Additionally a sensitivity analysis shows that the insignificance of education in previous studies is due to the test design conducted and cannot be interpreted as falsification of Borjas' and Bratsberg's (1996) theory.

JEL classification: J24, F22

Keywords: return migration, selective return migration

1. Introduction

Considering demographic change in Western European countries it is rather important to understand migration flows. Until now 70 million inhabitants in Europe have migrated internationally at least once (UC Davis (2010)). In order to evaluate the economic impact of migration for sending and receiving regions it is necessary to recognize the selection processes driving in- and outmigration. In particular understanding remigration is crucial to design and evaluate cohort related politics in the host countries concerned.

The thriving literature on this subject refers conceptually to the seminal paper by Borjas and Bratsberg (1996). Their model explicitly takes reversible migration into account. The authors describe the individual remigration decision primarily as a function of individual skills. They argue that the rate of return to skills in the source country relative to that in the host country determines the nature of selection. If the return to skills are relatively low (high) in the source country well (poorly) educated people will be prone to migrate to the host country. Selection in immigration is accentuated by complementary selection in outmigration. "The immigrants who remain in the host country are 'the best of the best' if there is positive selection, and the 'worst of the worst' if there is negative selection" (Borjas and Bratsberg (1996)).

The empirical literature referring to Borjas' and Bratsberg's (1996) theory concentrates on the analysis of remigration. Although there is a vast literature the empirical evidence on that issue remains nebulous and even the existence of the effect of skills on remigration is questionable thus far (Constant and Massey (2003)). On the one hand, using macro-level data of the United States Borjas and Bratsberg (1996) find supportive as well as contrary evidence for their model stated. On the other hand, there are many micro-level studies finding no significant effect of skills on return migration (Reagan and Olson (2000), Constant and Massey (2002), Constant and Massey (2003), Dustmann (2003), Kirdar (2009)).[1] An excellent synopsis of the literature on return migration can be found in Constant and Massey (2003).

Our inquiry is to shed light on the impact of skills on remigration flows by using German micro-level data. For this undertaking it is useful to distinguish a weak and a strong deduction from the seminal Borjas and Bratsberg (1996) model. The weak one states that the kind of the underlying selection process (positive or negative selection) determines the effect of skills on the remigration decision and, thus, is country-specific. The strong one additionally states the direction of that effect to be positive in the case of negative selection and negative in the case of positive selection.

We contribute to the existing literature in several aspects. Some of them exemplified below may explain the conflicting results of the aforementioned studies. Firstly, testing empirically more powerful hypotheses usually goes along with tougher data requirements. In the case of Borjas' and Bratsberg's

[1] The studies cited above differ regarding the estimation technique conducted and the definition of the endogenous variable as well as the measurement of skills or ability. Furthermore, there are many studies not mentioned analyzing remigration flows between two particular countries with mixed results as discussed in Constant and Massey (2003). The latter approach heavily depends on the assumption of time constant return to skills ratios for the countries considered.

theory (1996) (and depending on the level of aggregation) the measurement of country specific outmigration rates, the return to skills ratios between host and source countries over time, and the amount of transferable skills is especially challenging. By separating the theoretical insights of Borjas and Bratsberg (1996) in a weak and a strong hypothesis we clarify the empirical content of the theoretical statements under consideration and thus are able to test each of them with just the data requirements necessary. Secondly, our econometric analysis explicitly allows for endogeneity of host country specific capital as argued by DaVanzo (1983) and Dustmann (1996). We provide an in-depth discussion on that issue and estimate a so-called type II model which describes the return migration as a function of the actual amount of the country specific capital instead of the latent propensity to accumulate it. Finally, in order to ensure the robustness of our results and the comparability with previous studies we perform an extensive sensitivity analysis. We not only incorporate the most relevant formulation of the Borjas and Bratsberg (1996) model in previous research, but also use different kinds of endogenous variables discussed in the literature. In the course of this analysis we revisit the conflicting previous results and shed some light on their occurrence.

Based on the German Socio-Economic Panel (GSOEP) we find strong and robust evidence for the weak hypothesis of Borjas and Bratsberg (1996) formulated above. Additionally there is some evidence for the strong hypothesis as well. Moreover, most of the other factors supposed to determine return migration are in line with previous literature. Our results indicate that endogeneity has to be taken into account in order to avoid misspecification error if one relies on return intentions. Lastly, the sensitivity analysis performed in this study suggests that, according to Borjas and Bratsberg (1996), the theoretically unexpected insignificance of education on return migration is due to the common formulation of Borjas' and Bratsberg's (1996) model in the literature and should not be interpreted as an empirical falsification of their theory.

This paper is organized as follows. In the next section we describe our empirical implementation of the theory by Borjas and Bratsberg (1996) and address data measurement issues for both stages of our two stage approach in general. In the third section we present the data used in our study. Econometric issues are discussed in the fourth section of the paper. In section five the results are given and their sensitivity to alternative specifications is considered. A final section concludes.

2. Empirical Implementation

As mentioned above, to assess the impact of skills on the individual return migration decision we refer to Borjas and Bratsberg (1996) and distinguish a weak and a strong hypothesis. In their seminal paper they model the individual remigration probability and state that there is a selection process on skills, this process is country-specific, and the type of selection depends on the ratio of the return to skills between the source and the host country. To test the entire empirical proposition of this theory we need information about the ratio of return to skills for all source countries. These measures commonly suffer from substantial error. To avoid this potential error it seems useful to reduce the empirical content of the hypothesis tested. We therefore define the weak hypothesis to be:

(H1) The effect of skills on the return migration decision does vary over source countries.

This so-called weak hypothesis is nested in the strong one which additionally states the type of the selection under consideration. We therefore define:

(H2) If the return to skills in the source country are greater (smaller) than those in Germany the effect of skills on the return migration decision is positive (negative).

By using micro-level data we model the individual return migration decision (return_i) as follows. We observe $i = 1,\ldots,N$ individuals from $j = 1,\ldots,J$ source countries. According to Borjas' and Bratsberg's (1996) model Equation (1a) depicts the weak hypothesis and the more restrictive Equation (1b) the strong one, respectively.

$$\text{(1a)} \qquad \text{return}_i = \alpha + \sum_j D_j \cdot \delta_j + \text{skills}_i \cdot \varphi_W + \sum_j D_j \cdot \text{skills}_i \cdot \phi_{W,j} + C_i \cdot \gamma + I_i \cdot \mu + S_j \cdot \theta + \varepsilon_i,$$

$$\text{(1b)} \qquad \text{return}_i = \alpha + \sum_j D_j \cdot \delta_j + \text{skills}_i \cdot \varphi_S + T \cdot \text{skills}_i \cdot \phi_S + C_i \cdot \gamma + I_i \cdot \mu + S_j \cdot \theta + \upsilon_i.$$

Where α, δ_j, φ_W, φ_S, $\phi_{W,j}$, ϕ_S, γ, μ, and θ denote parameters indexed for the weak and the strong hypothesis respectively and υ_i and ε_i are i.i.d. error components. In both equations the selection process formulated by Borjas and Bratsberg (1996) is captured by the interaction terms involving individual skills (skills_i).

The interaction term between skills and country-specific dummy variables ($D_j \cdot \text{skills}_i$) in Equation (1a) formalizes the weak hypothesis and thus allows for country-specific skill effects. If the weak hypothesis (H1) by Borjas and Bratsberg is true[2] we will observe differences between the scalar parameter φ_W which is the skill effect for the base group, and the ones for individuals from other countries. The differences are measured by the parameters $\phi_{W,j}$. The joint test for all $\phi_{W,j}$ can be interpreted as testing the weak hypothesis (H1), where the null hypothesis is that it does not hold.

The empirically more ambitious strong hypothesis (H2) is tested in Equation (1b) where the variable *T* reflects the idea of different types of selection mechanism. *T* can be designed in multiple ways. For the time being let *T* separate source countries into two groups, distinguishing the countries with greater return to skills $(T = 1)$ from the ones with smaller return to skills $(T = 0)$ relative to Germany. Another, more ambitious, way to think about *T* is to assume that it measures the standardized difference between the return to skills in the host country $(r_{Germany})$ and the source country (r_{source}) metrically $\left(T = \left(r_{source} - r_{Germany}\right) / r_{Germany}\right)$ and, hence, is distributed around zero. In this case by interacting skills and *T* we implicitly assume a linear relationship between the impact of skills on remigration and *T*. According to Borjas and Bratsberg (1996) in either case we expect to find a positively signed scalar parameter ϕ_S. In what follows we use both concepts of *T*.

[2] It also is necessary that the host countries differ over the sample with respect to their return to skills.

All other possible source country specific selection processes are captured by the source country characteristics (S_j) and the level shift parameters δ_j .

To account for individual specific variables determining the return migration decision we include socio-demographic characteristics of two kinds in both equations. In line with the existing literature we choose a set of exogenous variables (I_i). Furthermore, we include variables measuring locally tied capital (C_i) in the model. These location specific assets can hardly be utilized after remigration and their liquidation induces high transaction cost. That is why we generally expect negative effects of host country specific capital on the return migration decision. At this point the problem of endogeneity arises. In general, we follow the reasoning in DaVanzo (1983) and Dustmann (1996) who conjecture that host country specific capital (C_i) is determined simultaneously with the individual return migration decisions, but there is one important distinction.

For reasoning simultaneity DaVanzo (1983) and Dustmann (1996) implicitly have to rely on the concept of proneness to accumulate such location specific capital. We, on the contrary, assume that the proneness to accumulate this kind of capital does not determine the outmigration decision, for the proneness is not region specific. For example, it is not the proneness to own a house that determines the return migration decision it is the fact that an individual owns a house or not. Hence, the intended return of individual *i* in period *t* is influenced only by the amount of social capital in period *t*. Although the proneness is correlated with that amount it alone does not affect anyone's remigration decision. In contrast to the amount, the proneness is mobile. To sum up, the proneness to remigrate in period *t-1* affects the proneness to accumulate location specific capital in period *t* as well as the proneness to remigrate in period *t*. So, we decide to model endogeneity rather than simultaneity. The corresponding estimation technique is a type II model which allows us to include endogenous dummy regressors rather than their reduced form predictions. A detailed discussion on this topic can be found in the section Econometric Issues.

3. Operationalizing the Variables

Our empirical analysis is based on a nationally representative data set for Germany, the German Socioeconomic Panel (GSOEP)[3]. It includes many relevant socio-economic variables and also oversamples immigrants, thus providing a large sample. The existing literature takes advantage of those features since a vast majority is based on GSOEP data. This fact also ensures the comparability of our results.

Measuring the endogenous variable, namely the return migration decision, is a well discussed issue in the existing literature. One strand of the existing empirical studies approximates the individual return migration decision modeled by Borjas and Bratsberg (1996) using the intended returns (Dustmann (1996), Steiner and Velling (1994)). Another strand argues that this approach can be improved by using the actual return moves (Constant and Massey (2003)). Employing the return intention is theoretically

[3] The data used in this publication were made available by the German Socio Economic Panel Study (SOEP) at the German Institute for Economic Research (DIW), Berlin. (Wagner, Frick and Schupp (2007)).

more appealing because human behavior is guided by this ontologically subjective category. The advantage of the latter concept is that solely actual return migration is relevant to politics. Yet, there are several shortcomings of the actual return concept. Since it is necessary to follow the same individuals for numerous years in order to figure out whether they return or not, panel attrition becomes a problem. Moreover one needs to assume that the ratio of return to skills and other country-specific characteristics remain unaltered (for a detailed discussion of this problem see Barrett and Trace (1998)). Since the GSOEP provides both measures we ensure comparability to central aforementioned studies and underpin the political relevance of our analysis by using both concepts.[4] Therefore, we do not use the latest GSOEP wave providing intended returns and rely on the 2000 survey instead.[5]

The following question measures return intentions.

Do you want to stay in Germany forever? [GSOEP Code qp133]

(yes/no)

We generate a dummy variable indicating individuals that answer no. The actual returns are generated as follows. In the GSOEP the reason for individual non-response is documented. We follow previous research and generate a dummy variable indicating a realized return within the time span from 2001 to 2007 using the non-response item *moved abroad*. Due to panel attrition the problem of sample selection arises as pointed out by Constant and Massey (2003). We address this problem below.

Another crucial issue for testing the hypotheses stated above is to measure the skills driving the selection process. Borjas' and Bratsberg's (1996) theoretical model refers to transferable skills only. Non-transferable skills get lost when migrating and, hence, have to be considered as migration costs.[6] In line with the existing literature we assume that formal education is transferable and neglect transferable skills beyond formal education. Thus, we measure skills as years of education.

[4] Since we use the GSOEP both approaches suffer from the fact that it is impossible to tell whether a former immigrant remigrates or moves to a third country as mentioned by DaVanzo (1983). The potential advantage of the return intention approach is weakened in the GSOEP because the interviewee is just asked whether he wants to stay in Germany forever or not. So, the existing studies as well as this inquiry assume that outmigrants are return migrants, which as Hunt (2004) points out is not a tough restriction.

[5] We perform a cross section analysis because the return intention is solely available in few waves. Thus, a classical panel data analysis is not feasible.

[6] A detailed discussion on that issue can be found in Chiswick and Miller (2009). They show that the individual probability of being undereducated or overeducated can mainly be explained by individual characteristics controlled for in our study like years since migration, marital status, and language proficiency. The possibly problematic issue of country specific depreciation rates for transferred skills seems to be weakened by their results.

As depicted in Equation (1a) testing the weak hypothesis relies on interactions between source country dummy variables and individual skills. We construct a total of thirteen source country dummies.[7] For testing the strong hypothesis provided in Equation (1b) we implement further constraints by imputing information about the type of the selection process. As mentioned above there are two ways to construct *T* for our undertaking. In both cases we have to proxy the ratio of return to skills between the source country and Germany. Ideally one could measure the return to transferable skills by the return to schooling as suggested by Rooth and Sareela (2007). Yet, besides that these data are hardly available for all source countries and the year 2000, a full set of restrictive assumptions is required just for measuring return to schooling (a brief discussion can be found in Björklund and Kjellström (2002)). That is why, we follow Borjas and Bratsberg (1996)[8] by making use of income inequality, concretely the Gini index, instead.[9] We generate a nominally scaled *T* indicating countries with higher return to skills than Germany and a metrically scaled *T* assumed to measure the standardized difference between the return to skills in the source country and Germany as defined in section 2.

Moreover, we take advantage of the existing literature that provides an elaborate pool of covariates. Starting with the endogenous location specific capital we control for *own house* and *German citizenship*. The variables are obtained as follows.

Do you live in the flat as the main tenant, subtenant, or owner? [GSOEP Code qh22]

(owner/tenant)

Is your nationality German? [GSOEP Code qp119]

(yes/no)

The rich SOEP data enables us to control for all major individual related covariates previously discussed in the literature denoted $\left(I_i\right)$ in the Equations (1a) and (1b). Finally, we enhance our micro level data by source country specific variables $\left(S_j\right)$. As suggested by Borjas and Bratsberg (1996) we augment the equations explaining the return migration decision by geographical distance to the source country, a

[7] All source country dummies consist of at least ten observations with at least one actual or intended return migrant. Since the number of missing observations varies with the endogenous variable used we make use different country dummies for each endogenous variable.

[8] Borjas and Bratsberg (1996) use the ratio of income accruing to the top 10% of the households to the income accruing to the bottom 20% of households to account for the ratio of return to skills between the host and the source country.

[9] The data are gathered from World Bank Development Indicators 2007. The time of their collection ranges from 1992 to 2004. However, 50% of the data refer to the time span of 2000 to 2003.

proxy for political stability, and GDP per capita.[10] Table 1 presents the definition and the summary statistics of all variables used in this inquiry.

Following Borjas and Bratsberg (1996) we restrict our sample by excluding all women from the sample to avoid the implicit assumption that the return migration process is the same for both sexes. Furthermore, we restrict our sample to all male immigrants of working age by our means between 18 and 65 years of age and not serving their community service or military service.

[10] We use the geographical distance from the source country's capital to Berlin, Germany. In addition, we proxy political stability using the sum of the political rights and the civil liberty indexes provided by the Freedom House Organization. We accessed the data from: http://www.freedomhouse.org/template.cfm?page=439, 11th September 2009. We obtain the real GDP per capita relative to the US from Heston, Summers and Aten (2009). For aggregated political entities like Benelux and Ex-Yugoslavia we calculate the population weighted average GDP/Capita, distance, and political stability proxy respectively. Furthermore, since realized returns refer to the time span from 2000 to 2007 we use the average value of all time varying source country specific variables denoted b) in Table 1.

Variable Definition			Summary Statistics			
Variable	Description	GSOEP	Mean	Sd.	Min;Max	Nobs.
ENDOGENOUS VARIABLES						
intended return	=1 if the respondent does not intend to stay in Germany	qp133	0.28	0.45	0;1	1169
realized return	=1 if the respondent actually outmigrates until 2007	ypergz$	0.10	0.30	0;1	897
own house	=1 if the respondent owns the flat or house he lives in	qh22	0.25	0.43	0;1	1404
citizen	=1 if the respondent is German citizen	qp119	0.39	0.49	0;1	1404
INDIVIDUAL CHARACTERISTICS (I_i)						
age	age in years	gebjahr	42.14	12.51	18;65	1404
age2	age squared	gebjahr	1932.88	1076.75	324;4225	1404
remitsum	remittances to the source country last year in Deutsche Mark	qp66*	579.59	2322.65	0;35000	1404
remitno	=1 if respondent is not remitting to the source country	qp66*	0.84	0.37	0;1	1404
healthgood	=1 if self reported health is above moderate	qp95	0.57	0.49	0;1	1404
healthmedi	=1 if self reported health is moderate (base outcome)	qp95	0.29	0.45	0;1	1404
healthbad	=1 if self reported health is below moderate	qp95	0.13	0.34	0;1	1404
ysm	years passed since initial migration to Germany	immiyear	19.53	10.97	0;51	1404
unemployed	=1 if respondent is not full or part-time employed	qp10	0.29	0.46	0;1	1404
partner_no	=1 if no partner exists	partz00	0.13	0.34	0;1	1404
partner_german	=1 if partner lives in the household and is from Germany	partz00	0.18	0.38	0;1	1404
partner_same	=1 if partner of same origin lives in the household (base)	partz00	0.56	0.50	0;1	1404
partner_different	=1 if partner lives in the household and is not of same origin	partz00	0.08	0.27	0;1	1404
partner_separated	=1 if partner not in the household and her origin is unknown	partz00	0.05	0.22	0;1	1404
education	years of schooling (Cross National Equivalent File)	d1110900	10.68	2.34	7;18	1404
prestige	Treiman standard international occupation prestige	siops00	27.83	18.94	0;78	1404
readgerman	=1 if respondent solely or usually reads German newspaper	qp129	0.56	0.50	0;1	1404
speakgerman	=1 if respondent solely or usually speaks German	qp128	0.40	0.49	0;1	1404
children	=1 if children live in the household	qkzahl	0.52	0.50	0;1	1404
SOURCE COUNTRY CHARACTERISTICS (S_j)						
GDP/capita[a),b)]	real gross domestic product per capita relative to the US	-	31.41[a)]	22.37[a)]	0.85;156.35[a)]	1404
	(a) indicates the value in 2000, (b) the average from 2000-	-	33.88[b)]	22.82[b)]	1.44;169.87[b)]	1404
rule of law[a),b)]	sum of Freedom House Indexes (political rights and civil	-	6.51[a)]	3.25[a)]	2;14[a)]	1404
	(a) indicates the value in 2000, (b) the sum from 2000-2007	-	39.11[b)]	22.39[b)]	14;98[b)]	1404
distance	distance in km from Berlin to source country's capital	-	1903.06	1611.05	282.2;16064	1404
T	stand. difference of gini-indexes ($T = (r_{source} - r_{Germany}) / r_{Germany}$)	-	0.29	0.21	-0.17;1.63	1350
SOURCE COUNTRY DUMMY VARIABLES (D_j)						
turkey	=1 if respondent was born in Turkey (base group)	corigin	0.25	0.43	0;1	1404
france	=1 if respondent was born in France	corigin	0.01	0.10	0;1	1404
great britain	=1 if respondent was born in Great Britain	corigin	0.01	0.11	0;1	1404
greece	=1 if respondent was born in Greece	corigin	0.05	0.22	0;1	1404
italy	=1 if respondent was born in Italy	corigin	0.10	0.30	0;1	1404
kasakhstan	=1 if respondent was born in Kazakhstan	corigin	0.07	0.25	0;1	1404
poland	=1 if respondent was born in Poland	corigin	0.11	0.31	0;1	1404
romania	=1 if respondent was born in Romania	corigin	0.03	0.17	0;1	1404
spain	=1 if respondent was born in Spain	corigin	0.02	0.15	0;1	1404
usa	=1 if respondent was born in the USA	corigin	0.01	0.10	0;1	1404
africa	=1 if respondent was born in Africa	corigin	0.02	0.14	0;1	1404
benelux	=1 if respondent was born in a BENELUX country	corigin	0.01	0.10	0;1	1404
ex-yugoslavia	=1 if respondent was born in Ex-Yugoslavia	corigin	0.12	0.32	0;1	1404
south america	=1 if respondent was born in South America	corigin	0.01	0.10	0;1	1404
other	=1 if respondent was born in non of the countries above	corigin	0.19	0.39	0;1	1404

Variable Definition			Summary Statistics			
Variable	Description	GSOEP	Mean	Sd.	Min;Max	Nobs.
INSTRUMENTAL VARIABLES FOR OWN HOUSE (EQUATION 2)						
married	=1 if respondent is married	qp140	0.78	0.42	0;1	1404
schooling_no	=1 if respondent has not graduated from school	qpsbil(a)	0.12	0.32	0;1	1404
schooling_hi	=1 if respondent has a high school degree	qpsbil(a)	0.17	0.38	0;1	1404
number of children	number of children in household	qkzahl	0.96	1.16	0;8	1404
car	=1 if respondent has access to a car for personal use	qp06	0.72	0.45	0;1	1404
wage	Average monthly net labor income in Euro in 2000	labnet00	1178.54	967.55	0;8180.64	1404
density	population density in respondents region of residence	-	582.12	635.08	56;3793	1404
INSTRUMENTAL VARIABLES FOR CITIZEN (EQUATION 3)						
marriedgerman	=1 if respondent has a German wife	qp140	0.15	0.36	0;1	1404
xenophobia	=1 if respondent worries about xenophobia in Germany	qp11809	0.30	0.46	0;1	1393
feelgood	self reported overall satisfaction with life	qp14301	7.02	1.80	0;10	1404
feeldisplaced	=1 if respondent feels alien both in host and home country	qp132	0.14	0.35	0;1	1178
feelbackhome	=1 if respondent's time to acclimatize in home country is short	qp137	0.50	0.50	0;1	1103
feeldiscriminated	=1 if respondent experienced discrimination due to origin	qp127	0.49	0.50	0;1	1180
aussiedler	=1 if respondent is an ethnic German	biimgrp00	0.10	0.31	0;1	1404
guest worker	=1 if respondent belongs to the 'guest worker subsample'	psample	0.39	0.49	0;1	1404
i90	=1 if respondent immigrated between 1980 and 1990	immiyear	0.18	0.39	0;1	1404
i00	=1 if respondent immigrated between 1990 and 2000	immiyear	0.39	0.49	0;1	1404

NOTE: *The summary statistics given above are based on the set union of all observations used in at least one of the estimated models presented in this paper. In order to save space the constant and dummy variables capturing missing values of remittances, years since migration, language proficiency, source country specific variables, and the self reported feelings on Germany are not reported (available upon request).*

4. Econometric Issues

Since ownership and German citizenship are supposed to be endogenous variables, we specify a three equation model that allows for endogeneity. All endogenous variables are binary. Thus, linear regression equations for these variables are not appropriate. As usual we define the binary variables as indicators for a latent tendency and explain the latent tendencies via linear regression equations. To be more precise, let y_1, y_2, and y_3 be the dummy indicators for (intended or realized) return, ownership, and German citizenship and y_1^*, y_2^*, and y_3^* the corresponding latent tendencies. Furthermore, all exogenous variables (including the constant) are collected in the vectors x_1, x_2, and x_3. Then we get:

$$(2) \qquad y_{1i} = \begin{cases} 1 & \text{if } y_{1i}^* > 0 \\ 0 & \text{else} \end{cases}, \quad y_{2i} = \begin{cases} 1 & \text{if } y_{2i}^* > 0 \\ 0 & \text{else} \end{cases}, \quad y_{3i} = \begin{cases} 1 & \text{if } y_{3i}^* > 0 \\ 0 & \text{else} \end{cases},$$

$$(3) \qquad \begin{aligned} y_{1i}^* &= \gamma_{12} y_{2i} + \gamma_{13} y_{3i} + \beta_1' x_{1i} + u_{1i} \\ y_{2i}^* &= \gamma_{23} y_{3i} + \beta_2' x_{2i} + u_{2i} \qquad \text{with } i = 1, \ldots, N. \\ y_{3i}^* &= \beta_3' x_{3i} + u_{3i} \end{aligned}$$

We partly use different exogenous variables in the three equations in order to strengthen the identification of the system (3).

In multiple equation models with endogenous binary variables either the endogenous binary variables themselves or the latent tendencies can be used as regressors (type II versus type I modeling, see Blundell and Smith (1993) for an overview). As discussed above, we argue that individual return migration decisions are rather determined by the existence of host country specific capital than the mere proneness to accumulate it. Therefore, a so-called type II specification with the actually observed binary variables on the right hand side is appropriate. Furthermore, since economic theory suggests a recursive model structure, as discussed above, the problem of logical inconsistency common to some type II specifications does not arise (see Wilde (2001) for a more detailed discussion of logical consistency and its implications for the model structure).

In model (3) endogeneity is equivalent to a non-zero correlation between the disturbances of the three equations. We will test this with a likelihood ratio test to clarify whether endogeneity is a statistically significant problem or not. Furthermore, we make use of the standard assumptions for the vector $u_i = (u_{1i}, u_{2i}, u_{3i})'$ of disturbances given the vector x_i of all exogenous variables, i.e.

$$(4) \qquad u_i \mid x_i \sim N(0, \Sigma), \quad \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{pmatrix}, \quad Cov(u_i, u_k \mid x_i) = 0 \quad \text{if } i \neq k, \text{ with } i,k = 1,\ldots,N$$

The assumption of uncorrelated vectors of disturbances of different individuals is not restrictive because we estimate a cross-section of our data set.

It is well known that the Maximum Likelihood Estimator (MLE) is asymptotically efficient. Since we use a large data set, we can make use of this efficiency. However, optimizing a (log)likelihood function based on a trivariate normal distribution is still computationally burdensome. Therefore, we use the Maximum Simulated Likelihood Estimator (MSLE), where the unknown probabilities are replaced by the mean of R simulated probabilities. If N, $R \to \infty$ and $\sqrt{N}/R \to 0$, the MSLE is asymptotically equivalent to the MLE. [11] Different methods for the simulation have been proposed in the literature. The most accurate one seems to be the GHK simulator (Greene (2008)), which we use. A crucial point is choosing R. Although Börsch-Supan and Hajivassiliou (1993) found in the simulation studies of their seminal paper that even R = 20 leads to a negligible bias, we need a higher R to get stable results. Additionally, a higher R involves higher efficiency. We choose R = 10000.

Concerning the realized returns panel attrition may be an issue. This is usually dealt with by using multinomial models (e.g. Constant and Massey (2003)). However, until now a combination of type II models and multinomial models does not exist. Therefore we estimate both types of models for the realized returns.

[11] Cameron and Trivedi (2005), chap. 12.4.2, the last condition is only needed for efficiency and not for consistency.

5. Results and Sensitivity Analysis

We estimate the models specified in Equation (1a) and (1b) allowing for endogeneity as discussed in section 2 and 4.[12] We perform a number of Wald-Tests, given at the bottom of Table 2, indicating the high statistical relevance of all specifications considered in this study. Generally our results show that the occurrence of the problem of endogeneity depends on the dependent variable used. Endogeneity only becomes an issue if we make use of return intentions on the left hand side as can be seen from the Wald-Tests at the bottom of Table 2 referring to the Null-Hypothesis of no correlation between the residuals of different equations in system (3).[13] The same Wald-Tests performed with actual returns on the left hand side of the system (3) never resulted in a p-value below 0.2769.[14] This could reflect the notably theoretical difference between both measurement concepts. Since it refers to the following years, actual remigration is less determined by factors related to the amount of country specific capital in period *t* not controlled for in the model than the return intention in this period. Therefore, considering realized returns, estimating an ordinary multinomial logit model is superior. In what follows regarding realized returns we refer to the results of a multinomial logit model whereas the results regarding intended returns stem from the recursive three equation type II probit model discussed in chapter 4.[15]

To begin with we refer to the first four columns of Table 2 where we test the weak and the strong hypothesis by Borjas and Bratsberg (1996) using intended returns as well as the actual return decision. At first let us turn to the weak hypothesis stated above and formalized in Equation (1a).

The weak hypothesis

As pointed out in section 2 the weak hypothesis by Borjas and Bratsberg (1996) states that the effect of skills on return migration is determined by the type of selection process at hand and therefore is country specific. The parameter estimates for the weak hypothesis (H1) formalized in Equation (1a) for both endogeneous variables and estimation techniques are given in the first two columns of Table 2. In order to test this hypothesis we conduct a Likelihood-Ratio Test which tests the Null Hypothesis of no differences in skill effects between source countries. That is, all parameters $\phi_{W,j}$ equal zero. We can reject the Null for intended returns and actual returns with p-values of 0.0016 and 0.0000, respectively. These results strongly support the weak Borjas and Bratsberg (1996) hypothesis.

The strong hypothesis

The strong hypothesis (H2) formalized in Equation (1b) adds empirical content to the weak one by stating the direction of the impact of skills on remigration depending on the type of the selection process observed. Altogether we estimate four models using both concepts of *T* indicating the type of selection

[12] We make use of the STATA routine 'triprobit' written by Terracol (2002). The reported results are based on robust standard errors.

[13] According to Equation (4) the Null Hypothesis can be formalized as: $H_0: \sigma_{12} = \sigma_{13} = \sigma_{23} = 0$.

[14] The entire results of all estimations not reported in the paper are available from the authors upon request.

[15] The parameter estimates for the unspecified dropout alternative in the case of the multinomial logit model as well as the estimation results for the second and third equation of the recursive three-equation probit model are given in the appendix.

process at hand for both endogenous variables considered in our inquiry. Since both concepts of *T* show the same results we focus on the results for the more ambitious metrically one. The parameter estimates for the strong hypothesis are presented in the third and fourth column of Table 2. According to Borjas' and Bratsberg's (1996) theory we expect to find a positively signed parameter for the interaction of skills and *T*. For realized returns the results show the expected sign. Hence, the higher the return to skills in the source country relative to Germany the larger is the impact of skills on remigration. In the case of negative selection higher education fosters outmigration whereas in the case of positive selection higher educated individuals are discouraged to remigrate. This finding is significant at the five percent level. Concerning the intended returns no significance can be shown.

Control Variables

Regarding the control variables our findings are in line with previous literature and as theoretically expected. Especially the estimated effects of individual investment in host country attached capital on return intentions are significantly negative. The insignificance of locally tied capital on realized returns seems to reflect the fact that during a period of seven years the mobility discouraging effect of that investment is weakened. Nonetheless the sign is as expected.

One result deserves closer attention. According to Borjas and Bratsberg (1996) we expect to find a negative impact of the distance to the immigrant's home country. To the contrary we find a positive effect of distance on return intentions. We suggest a broader approach to think about geographical distance. Besides the cost of moving the geographical distance may reflect the cultural distance and, hence, may capture the unobserved cost of living in a culturally more distinct society. In the long run, associated with the realized return measure, both costs seem to outweigh each other resulting in no significant effect. In the short run, instead, a kind of homesickness may be more relevant causing the significant positive effect.

Comparison with previous research and robustness check

In this subsection of our paper we discuss the sensitivity of our results with respect to different model specifications and relate them to previous literature. First of all, the empirical evidence in favor of the weak hypothesis stated above proves to be robust regarding the different endogenous variables used. For the strong hypothesis the empirical evidence depends on the endogenous variable considered. Nonetheless we find some evidence supporting the strong hypothesis by Borjas and Bratsberg (1996). In general, our main findings remain stable across different endogenous variables and estimation techniques.[16]

As mentioned above using different estimation methods as well as endogenous variables many aforementioned micro-level studies do not find a significant effect of skills on remigration. This seems to contradict the importance attached by Borjas and Bratsberg (1996) to the impact of skills on outmigration. In order to shed some light on this issue we reestimate our models specifying the effect of

[16] Furthermore, since some observations may be classified as outliers we reestimated all models without them. All results presented below turn out to be robust to this kind of problem.

skills on remigration in the way commonly done in previous literature. As can be seen in the fifth and the sixth column of Table 2 the insignificance of education seems to be induced by this restrictive way in which the impact of skills is commonly formalized. Since the impact of skills depends on the type of selection process at hand we do not observe an overall skill effect on outmigration.

Table 2: Estimation Results

ENDOGENOUS VARIABLE	Weak Hypothesis				Strong Hypothesis				Previous Literature			
	Intended Returns		Realized Returns		Intended Returns		Realized Returns		Intended Returns		Realized Returns	
	coefficient	p-value	coefficient	p-value	coefficient	p-value	coefficient	p-value	coefficient	p-value	coefficient	p-value
COUNTRY-SKILL INTERACTIONS												
education	-0.0305	0.464	0.4099	0.002	0.0009	0.974	-0.1453	0.059	0.0116	0.644	-0.0639	0.335
educ *x* france (base: turkey)	0.1748	0.420	-0.0349	0.940	-	-	-	-	-	-	-	-
educ *x* great britain	0.0268	0.785	8.9952	0.981	-	-	-	-	-	-	-	-
educ *x* greece	-0.0518	0.471	-0.7396	0.005	-	-	-	-	-	-	-	-
educ *x* italy	-0.1164	0.132	-0.7916	0.001	-	-	-	-	-	-	-	-
educ *x* kasakhstan	-0.3099	0.020	-	-	-	-	-	-	-	-	-	-
educ *x* poland	0.0856	0.394	-0.2569	0.444	-	-	-	-	-	-	-	-
educ *x* romania	-0.0940	0.614	0.2860	0.567	-	-	-	-	-	-	-	-
educ *x* spain	-0.0869	0.493	-0.3510	0.246	-	-	-	-	-	-	-	-
educ *x* usa	0.5560	0.012	-1.0412	0.019	-	-	-	-	-	-	-	-
educ *x* africa	-0.5237	0.011	-0.6255	0.294	-	-	-	-	-	-	-	-
educ *x* benelux	0.0991	0.583	7.2662	0.981	-	-	-	-	-	-	-	-
educ *x* ex-yugoslavia	0.1204	0.065	-0.9050	0.000	-	-	-	-	-	-	-	-
educ *x* south america	-	-	-0.5705	0.154	-	-	-	-	-	-	-	-
educ *x* other	0.1584	0.012	-0.4797	0.017	-	-	-	-	-	-	-	-
T *x* educ	-	-	-	-	0.0307	0.409	0.3586	0.015	-	-	-	-
ENDOGENOUS REGRESSORS												
citizen	-1.0069	0.043	-0.9644	0.167	-1.0358	0.029	-1.0519	0.125	-0.9985	0.029	-1.0390	0.114
own house	-1.0968	0.040	-0.4823	0.215	-1.0803	0.067	-0.3556	0.329	-1.1130	0.043	-0.3681	0.311
COUNTRY-SPECIFIC CONTROLS												
GNP/capita	0.0062	0.339	0.0264	0.195	0.0092	0.222	0.0292	0.166	0.0084	0.205	0.0019	0.901
rule of law	-0.0141	0.739	-0.0682	0.010	-0.0122	0.784	-0.0672	0.007	-0.0154	0.707	-0.0826	0.001
distance	0.0001	0.009	0.0002	0.173	0.0001	0.038	0.0000	0.886	0.0001	0.007	0.0002	0.188
INDIVIUAL-SPECIFIC CONTROLS												
age	-0.0260	0.360	-0.0412	0.637	-0.0306	0.284	-0.0525	0.532	-0.0309	0.274	-0.0458	0.579
age2	0.0003	0.429	0.0007	0.465	0.0004	0.286	0.0008	0.397	0.0004	0.276	0.0007	0.435
remitsum	0.0001	0.007	0.0000	0.515	0.0001	0.010	0.0001	0.295	0.0001	0.010	0.0001	0.312
remitno	-0.0586	0.684	-0.5211	0.288	-0.0950	0.515	-0.1326	0.771	-0.0955	0.510	-0.0618	0.891
healthgood	-0.2523	0.010	0.5385	0.106	-0.2030	0.039	0.5442	0.087	-0.2037	0.038	0.5308	0.091
healthbad	-0.1879	0.195	-0.4136	0.376	-0.1280	0.359	-0.3397	0.455	-0.1273	0.360	-0.2809	0.530
ysm	0.0057	0.565	-0.0572	0.003	0.0055	0.555	-0.0561	0.002	0.0052	0.566	-0.0551	0.002
children	0.1041	0.366	0.1053	0.756	0.1482	0.182	0.1568	0.633	0.1495	0.175	0.0943	0.771
unemployed	0.0185	0.912	0.9222	0.080	0.0873	0.604	1.0705	0.029	0.0799	0.630	1.0959	0.025

Table 2: (continued)

	Weak Hypothesis				Strong Hypothesis				Previous Literature			
ENDOGENOUS VARIABLE	Intended Returns		Realized Returns		Intended Returns		Realized Returns		Intended Returns		Realized Returns	
	coefficient	p-value	coefficient	p-value	coefficient	p-value	coefficient	p-value	coefficient	p-value	coefficient	p-value
partner_german	0.1145	0.385	-1.6898	0.001	0.0629	0.624	-1.6938	0.001	0.0611	0.632	-1.6182	0.001
partner_no	0.1427	0.385	0.6293	0.164	0.1380	0.402	0.5386	0.214	0.1377	0.401	0.4761	0.266
partner_different	0.2196	0.174	-0.0050	0.993	0.2127	0.203	0.0176	0.973	0.2094	0.207	0.1905	0.700
partner_separated	-0.0063	0.979	-0.4019	0.590	0.0444	0.846	-0.1750	0.792	0.0484	0.832	0.0658	0.919
prestige	0.0060	0.145	-0.0058	0.661	0.0074	0.062	0.0035	0.771	0.0073	0.064	0.0065	0.588
readgerman	-0.3825	0.000	-0.5882	0.096	-0.3974	0.000	-0.7182	0.036	-0.3941	0.000	-0.7951	0.019
speakgerman	-0.2152	0.054	-0.4770	0.237	-0.2354	0.033	-0.5198	0.178	-0.2324	0.035	-0.4506	0.237
SOURCE-COUNTRY DUMMIES												
france (base: turkey)	-1.5541	0.531	-0.2550	0.964	0.4444	0.344	1.2710	0.291	0.3574	0.424	0.8966	0.439
great britain	-0.3218	0.794	-151.1389	0.980	-0.1765	0.675	-1.7295	0.239	-0.2239	0.580	-1.4734	0.305
greece	1.0065	0.191	6.7026	0.012	0.4673	0.091	0.4613	0.580	0.3818	0.117	0.1241	0.869
italy	1.3074	0.106	6.7656	0.007	0.1559	0.595	-0.3854	0.666	0.1084	0.688	-0.2970	0.727
kasakhstan	2.2044	0.145	-	-	-0.6740	0.175	-	-	-0.8255	0.063	-	-
poland	-0.8560	0.499	-0.0330	0.994	0.1895	0.616	-1.3405	0.265	0.0714	0.837	-2.3166	0.046
romania	1.2853	0.570	-4.7337	0.495	0.3213	0.511	1.7468	0.162	0.1457	0.739	0.0182	0.987
spain	1.3040	0.297	2.4555	0.425	0.4913	0.160	-0.2707	0.797	0.4148	0.198	-0.4452	0.655
usa	-7.4734	0.009	10.2794	0.086	-0.3141	0.646	-1.9337	0.278	-0.3618	0.571	-0.9958	0.509
africa	5.5791	0.008	7.3628	0.306	0.2921	0.419	-1.8972	0.472	0.2171	0.520	0.3962	0.788
benelux	-0.7836	0.752	-132.3133	0.981	0.3423	0.545	0.2528	0.872	0.2463	0.652	-0.9867	0.581
ex-yugoslavia	-0.9273	0.160	9.3803	0.000	0.4091	0.099	2.1346	0.009	0.2629	0.139	0.5691	0.295
south america	-	-	6.3157	0.216	-	-	-0.3413	0.847	-	-	0.4757	0.746
other	-1.6204	0.026	4.0539	0.089	0.1797	0.490	0.4470	0.605	0.0818	0.713	-0.4390	0.571
RESIDUAL CORRELATION COEFF.												
rho12	0.5493	0.043	-	-	0.470	0.064	-	-	0.4695	0.054	-	-
rho13	0.4698	0.179	-	-	0.461	0.236	-	-	0.4838	0.185	-	-
rho23	0.0422	0.798	-	-	0.063	0.696	-	-	0.0633	0.693	-	-
	test statistic	p-value	teststatistic	p-value	teststatistic	p-value	teststatistic	p-value	teststatistic	p-value	teststatistic	p-value
Wald Test: *Residual Corr.*	11.27 (3)	0.010	-	-	9.26 (3)	0.026	-	-	9.65 (3)	0.022	-	-
LR-Test: *educ x country*	33.13 (13)	0.001	49.74 (13)	0.000	-	-	-	-	-	-	-	-
OVERALL SIGNIFICANCE	803.72 (78)	0.000	387.39 (106)	0.000	710.36 (65)	0.000	349.27 (80)	0.000	712.00 (65)	0.000	337.64 (80)	0.000
NUMBER OF OBSERVATIONS	1166		1383		1166		1383		1166		1383	

NOTE: *IN THE CASE OF INTENDED RETURNS THE ESTIMATION RESULTS FOR EQUATION TWO AND THREE ARE GIVEN IN TABLE 3 IN THE APPENDIX, THE RESULTS FOR ALTERNATIVE 2 OF THE MULTINOMIAL LOGIT ESTIMATION ARE PRESENTED IN TABLE 4 IN THE APPENDIX, COUNTRY BASE GROUP IS TURKEY, THE CONSTANT AND DUMMY VARIABLES INDICATING MISSING OBSERVATIONS FOR REMITTANCES, YSM, SPEAKGERMAN, READGERMAN, CONTEXT DATA AND THE INTERACTION TERM OF EDUCATION AND MISSING GINI-COEFFICIENTS ARE NOT REPORTED*

6. Conclusion

More than ten years after the seminal paper by Borjas and Bratsberg (1996) modeling the impact of skills on remigration the empirical evidence on that theory is still mixed. This paper sheds light on that issue. We deduct two hypotheses from Borjas and Bratsberg (1996) theory, a weak one and a strong one. The weak hypothesis states that the skill effect on remigration is country specific whereas the strong hypothesis additionally argues that the impact of skills on remigration depends on the ratio of return to skills in the source country relative to the host country. Using the GSOEP supplemented with several context variables we test the hypotheses by Borjas and Bratsberg (1996) separately while allowing for endogeneity of host country specific capital. To our knowledge this is the first study formalizing and testing the hypotheses stated by Borjas and Bratsberg (1996) in this direct way using micro-level data.

In general, our findings give strong support for the theory by Borjas and Bratsberg (1996). Our main results prove to be robust to the different endogenous variables and estimation techniques employed. Yet, on closer inspection our inquiry also shows the remarkable differences resulting from the quite distinct theoretical measurement concepts connected with intended and realized returns. For example, the problem of endogeneity, the impact of distance between the host and source country, the effects of host country specific capital as well as the political stability depend on the endogenous variable used.

Finally, a sensitivity analysis shows that the insignificance of education reported in previous studies is due to the negligence of country specific skill effects in the empirical tests commonly conducted and should not be interpreted as falsification of Borjas' and Bratsberg's (1996) theory.

References

Barrett A, Trace F (1998) Who is coming back? The educational profile of returning migrants in the 1990s. Irish Banking Review (Summer): 38-52.

Björklund A, Kjellström C (2002) Estimating the return to investments in education: how useful is the standard Mincer equation? Economics of Education Review 21: 195–210.

Blundell R, Smith RJ (1993) Simultaneous microeconometric models with censored or qualitative dependent variables. In: Maddala GS, Rao CR, Vinod HD (ed) Handbook of statistics 11. North Holland, Amsterdam 117-143.

Borjas GJ, Bratsberg B (1996) Who leaves? The outmigration of the foreign-born. The Review of Economics and Statistics 78(1): 165-176.

Börsch-Supan A, Hajivassiliou VA (1993) Smooth unbiased multivariate probability simulators for maximum likelihood estimation of limited dependent variable models. Journal of Econometrics 58(3): 347-368.

Cameron AC, Trivedi PK (2005) Microeconometrics – Methods and applications. Cambridge University Press, Cambridge.

Chiswick BR, Miller PW (2009) The international transferability of immigrants' human capital. Economics of Education Review 28(2): 162-169.

Constant A, Massey DS (2002) Return Migration by German Guestworkers: Neoclassical versus New Economic Theories. International Migration 40(4): 5-38.

Constant A, Massey DS (2003) Self-selection, earnings, and out-migration: A longitudinal study of immigrants to Germany. Journal of Population Economics 16(4): 631-653.

DaVanzo J (1983) Repeat Migration in the United States: Who moves back and who moves on? The Review of Economics and Statistics 65(4): 552-559.

Dustman C, Bentolila S, Faini R (1996) Return Migration: The European Experience. Economic Policy 11(22): 213-250.

Dustmann C (2003) Children and Return Migration. Journal of Population Economics 16(4): 815-830.

Greene WH (2008) Econometric analysis. 6th edition, Pearson Education, Upper Saddle River (NJ).

Heston A, Summers R, Aten B (2009) Penn World Table Version 6.3, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania

Hunt J (2004) Are migrants more skilled than non-migrants? Repeat, return and same-employer migrants. Canadian Journal of Economics 37(4): 830-849.

Kirdar MG (2009) Labor Market Outcomes, Savings Accumulation, and Return Migration. Labour Economics 16(4): 418-428.

Reagan PB, Olson RJ (2000) You can go home again: Evidence from Longitudinal Data. Demography 37(3): 339-350.

Rooth DO, Saarela J (2007) Selection in migration and return migration: Evidence from micro data. Economics Letters 94: 90-95.

Steiner V, Velling J (1994) Re-Migration Behaviour and Expected Duration of Stay of Guest Workers in Germany. In: Steinmann G (ed) The Economic Consequences of Immigration to Germany. Physika, Heidelberg 101-119.

Terracol A (2002) TRIPROBIT: Stata module to estimate trivariate probit model using the GHK simulator, Statistical Software Components S424302, Boston College Department of Economics.

University of California (2010) Global: UN Migrants, Population. Migration News 17(1), online available: http://migration.ucdavis.edu/MN/comments.php?id=3585_0_5_0.

Wagner G, Frick J, Schupp J (2007) The German Socio-Economic Panel Study (SOEP) – Scope, Evolution and Enhancements. In: Schmollers Jahrbuch (Journal of Applied Social Science Studies) 127(1): 139-169.

Wilde J (2001) Simultaneity in multiple equation hybrid models with endogenous dummy regressors. Journal of the German Statistical Society 85: 267-281.

World Bank (2007) World Development Indicators, Development Data Group, Washington.

Appendix

Table 3: Estimation Results for Equations (2) and (3)

	Weak Hypothesis		Strong Hypothesis		Previous Literature	
ENDOGENOUS VARIABLE	Own House (Equation 2)					
	coefficient	p-value	coefficient	p-value	coefficient	p-value
ENDOGENOUS REGRESSORS						
citizen	0.2220	0.147	0.2045	0.178	0.2041	0.177
CONTROL VARIABLES						
age	0.0173	0.000	0.0174	0.000	0.0174	0.000
INSTRUMENTS						
married	-0.0435	0.752	-0.0431	0.758	-0.0412	0.767
schooling_no	-0.3077	0.047	-0.3321	0.052	-0.3350	0.046
schooling_hi	0.4258	0.001	0.4315	0.001	0.4326	0.001
density	-0.0002	0.077	-0.0002	0.068	-0.0002	0.069
number of children	0.1403	0.001	0.1401	0.001	0.1399	0.001
car	0.3689	0.001	0.3697	0.001	0.3698	0.001
wage	0.0002	0.000	0.0002	0.002	0.0002	0.002
ENDOGENOUS VARIABLE	Citizen (Equation 3)					
	coefficient	p-value	coefficient	p-value	coefficient	p-value
CONTROL VARIABLES						
unemployed	0.2572	0.073	0.2509	0.085	0.2509	0.085
rule of law	0.0902	0.000	0.0933	0.000	0.0932	0.000
INSTRUMENTS						
marriedgerman	-0.6661	0.004	-0.6621	0.006	-0.6619	0.006
xenophobia	0.0446	0.742	0.0386	0.781	0.0383	0.782
feelgood	0.1591	0.000	0.1568	0.000	0.1574	0.000
feeldisplaced	-0.6209	0.001	-0.6227	0.002	-0.6231	0.002
feelbackhome	-0.4103	0.003	-0.3893	0.005	-0.3886	0.005
feeldiscriminated	-0.0323	0.792	-0.0404	0.744	-0.0375	0.762
aussiedler	2.8761	0.000	2.8732	0.000	2.8699	0.000
gastarbeiter	-1.2089	0.000	-1.2403	0.000	-1.2376	0.000
i90	2.7234	0.000	2.7391	0.000	2.7406	0.000
i00	2.7335	0.000	2.7452	0.000	2.7457	0.000
NUMBER OF OBSERVATIONS			1166			

NOTE: *THE CONSTANT AND DUMMY VARIABLES INDICATING MISSING OBSERVATIONS FOR YSM, SELF REPORTED FEELINGS ON GERMANY, AND AUSSIEDLER ARE NOT REPORTED (AVAILABLE UPON REQUEST).*

Table 4: Estimation Results for the Multinomial Logit Model (alternative 2: unspecified dropout)

	Weak Hypothesis		Strong Hypothesis		Previous Literature	
ENDOGENOUS VARIABLE	Realized Returns (alternative 2: unspecified dropout)					
	coefficient	p-value	coefficient	p-value	coefficient	p-value
INDIVIUAL-SPECIFIC CONTROLS						
age	0.0195	0.653	0.0162	0.708	0.0181	0.674
age2	-0.0004	0.459	-0.0004	0.488	-0.0004	0.457
remitsum	0.0000	0.121	0.0001	0.122	0.0000	0.127
remitno	-0.1679	0.416	-0.1486	0.473	-0.1529	0.460
healthgood	-0.0304	0.839	-0.0402	0.786	-0.0374	0.800
healthbad	-0.1526	0.478	-0.1325	0.535	-0.1346	0.528
ysm	0.0118	0.209	0.0127	0.174	0.0128	0.168
employment	-0.0486	0.842	-0.0482	0.842	-0.0599	0.805
partner_german	-0.1454	0.470	-0.1613	0.416	-0.1604	0.418
partner_different	0.0911	0.717	0.0379	0.878	0.0429	0.862
partner_separated	0.3327	0.295	0.3015	0.340	0.3203	0.310
partner_no	0.2351	0.320	0.2190	0.351	0.2146	0.360
prestige	-0.0013	0.823	-0.0007	0.905	-0.0010	0.866
readgerman	-0.1529	0.368	-0.1451	0.387	-0.1409	0.400
speakgerman	-0.0567	0.719	-0.0647	0.680	-0.0597	0.703
children	-0.1227	0.435	-0.1283	0.410	-0.1374	0.377
citizen	-0.1982	0.401	-0.2054	0.384	-0.2170	0.349
property	0.3144	0.041	0.3183	0.038	0.3104	0.043
COUNTRY-SPECIFIC CONTROLS						
GNP/capita	0.0052	0.570	0.0038	0.658	0.0035	0.688
rule of law	0.0046	0.535	0.0019	0.802	0.0040	0.586
distance	0.0001	0.271	0.0001	0.441	0.0001	0.329
COUNTRY-SKILL INTERACTIONS						
education	0.0243	0.703	-0.0507	0.159	-0.0367	0.264
educ *x* france	1.4455	0.805	-	-	-	-
educ *x* great britain	-0.8017	0.829	-	-	-	-
educ *x* greece	0.6287	0.685	-	-	-	-
educ *x* italy	1.7046	0.224	-	-	-	-
educ *x* poland	1.4815	0.239	-	-	-	-
educ *x* romania	-0.3132	0.894	-	-	-	-
educ *x* spain	-2.7682	0.250	-	-	-	-
educ *x* usa	-2.5209	0.608	-	-	-	-
educ *x* africa	3.4698	0.110	-	-	-	-
educ *x* benelux	2.1485	0.656	-	-	-	-
educ *x* ex-yugoslavia	1.4348	0.225	-	-	-	-
educ *x* south america	0.0109	0.998	-	-	-	-
educ *x* other	0.5685	0.554	-	-	-	-
T *x* educ	-	-	0.0528	0.323	-	-
SOURCE-COUNTRY DUMMIES						
france	-0.1528	0.766	0.0679	0.933	-0.0527	0.947
great britain	0.0217	0.943	-0.2769	0.712	-0.3177	0.671
greece	-0.1197	0.417	-0.4163	0.366	-0.5239	0.237
italy	-0.1862	0.174	-0.0180	0.966	-0.0670	0.873

Table 4: (continued)

	Weak Hypothesis		Strong Hypothesis		Previous Literature	
ENDOGENOUS VARIABLE	Realized Returns (alternative 2: unspecified dropout)					
	coefficient	p-value	coefficient	p-value	coefficient	p-value
poland	-0.0492	0.643	1.0948	0.004	1.0155	0.005
romania	0.0375	0.846	0.4646	0.357	0.2694	0.557
spain	0.2082	0.372	-0.5906	0.326	-0.6767	0.254
usa	0.0831	0.798	-0.8710	0.393	-0.8532	0.406
africa	-0.2328	0.181	0.8405	0.126	0.7816	0.139
benelux	-0.2469	0.518	-0.7227	0.497	-0.7294	0.463
ex-yugoslavia	-0.1618	0.158	0.0324	0.925	-0.2041	0.405
south america	-0.0760	0.811	-0.7754	0.434	-0.6595	0.493
other	-0.0453	0.595	0.3825	0.293	0.1993	0.516
NUMBER OF OBSERVATIONS			1383			

NOTE: *THE CONSTANT AND DUMMY VARIABLES INDICATING MISSING OBSERVATIONS FOR REMITTANCES, YSM, SPEAKGERMAN, READGERMAN, AND CONTEXT DATA ARE NOT REPORTED (AVAILABLE UPON REQUEST).*