ORIGINAL PAPER

How important is homeland education for refugees' economic position in The Netherlands?

Joop Hartog · Aslan Zorlu

Abstract We use data on refugees admitted to The Netherlands that include registration of education in their homeland by immigration officers. Such data are seldom available. We investigate the quality and reliability of the registrations and then use them to assess effects on refugees' economic position during the first 5 years after arrival. The most remarkable finding is the absence of returns to higher education.

Responsible editor: Klaus F. Zimmermann

J. Hartog
School of Economics, University of Amsterdam, Amsterdam, The Netherlands

J. Hartog
IZA, Bonn, Germany

J. Hartog
Tinbergen Institute, Amsterdam, The Netherlands

J. Hartog (✉) · A. Zorlu
Amsterdam Institute for Advanced Labour Studies (AIAS), University of Amsterdam, Amsterdam, The Netherlands
e-mail: j.hartog@uva.nl

A. Zorlu
Department of Human Geography, University of Amsterdam, Amsterdam, The Netherlands
e-mail: a.zorlu@uva.nl

 ⓍSpringer

1 Introduction

Education is commonly held to be a key variable for an individual's economic success. Similarly, immigrants may benefit directly from their education, as they would in their homeland or like natives in their destination country. On top of that, more educated immigrants may be more effective in the transition to a new society and thus may benefit relatively more from the education they bring from their homeland and from any additional education they may acquire in their destination country (Chiswick and Miller 1994, 2003; Chiswick et al. 2005; Bratsberg and Ragan 2002). Thus, education is a core variable in analyses of immigrants' economic success, and it is an important component in determining the differences between natives and immigrants (Card 1999; Altonji and Blank 1999).

The literature on immigrants in Europe does not provide straightforward and unequivocal support for the hypothesis that immigrants' education is rewarded in the same monotonic relationship that holds for natives. Although returns to education are usually lower for immigrants than for natives, the patterns found are sometimes quite irregular. For example, Bevelander and Nielsen (2001) reported that for Yugoslavian immigrants, the effect of education is the same as for Swedes, no matter where they acquired it, whereas Nordic immigrants have less benefits from their education if they acquired it in Sweden. Husted et al. (2001) analysed the position of immigrants in Denmark, using six levels of education. For natives, the first four levels beyond compulsory add some 10–15% to the wage rate, the highest level adds some 30%. For refugees, the first five levels add nothing or even depress wages, and levels six and seven add some 15%. For non-refugees, the first three levels add nothing, the fourth depresses wages by 10%, the fifth adds 5% and the sixth adds 10 to 15%.

Several studies have indicated that it matters very much whether an immigrant's education has been acquired in the origin country before migration or in the destination country. Friedberg (2000) even showed that properly accounting for education obtained before migration can explain the initial earnings disadvantage of immigrants. Existing studies that make the distinction never have direct observations on the decomposition: It is always inferred, usually from highest level of education attained and age at immigration (Friedberg 2000; Nekby 2002; Cortes 2004; an exception is Kee (1993) who used direct observations for immigrants to The Netherlands). The lack of a robust standard pattern of returns to education for immigrants may well be related to a problematic measurement of their education. However, it might just as well be a real phenomenon, given the substantial heterogeneity by source country, motive and ease of entry in the destination country's economy.

In this paper, we shed some light on the issue by focusing on the benefits of homeland education for refugees. We use registration by immigration officers obtained when immigrants apply for admission to The Netherlands. We investigate the quality of the data and then use the observations to assess the importance for economic success during the first 5–6 years after admission. Our key finding is that education beyond secondary does not yield any additional monetary returns. After thoroughly testing for the reliability of this conclusion, we are convinced that this is a real effect.

In the next section we introduce our data. In Section 3, we discuss reliability of registered education. Section 4 presents analyses of the effect of education on the probability of employment and the probability of receiving a social benefit. Section 5 presents the analyses of the schooling effect on wages. Section 6 considers the relation between initial disadvantage and later annual growth rate. Section 7 concludes.

2 Data selection

2.1 The files

All immigration by non-Dutch citizens is registered in the Central Register Foreigners (*Centraal Register Vreemdelingen*, CRV), using information from the Immigration Police (*Vreemdelingen Politie*) and the Immigration and Naturalisation Service (*Immigratie—en Naturalisatie Dienst*, IND). The CRV register records immigration motive, and this allows identification of refugees. At our request, CBS, the Dutch Central Bureau of Statistics, has linked the data to the Municipal Register of Population (*Gemeentelijke Basisadministratie*, GBA). The GBA/CRV Register includes all non-Dutch immigrants who legally entered The Netherlands during 1990–2001 and registered in the population register, except those who have returned before January 1, 1998, those who naturalised to Dutch citizenship and those who have died.

It should be noted that the time of registration in the Municipal Register of Population (GBA) significantly differs from the time of registration in the Central Register Foreigners (CRV) because of possible illegal stay and the asylum application procedure that can take up to a few years. As the register takes stock every year on January 1, all immigrants leaving within the calendar year of arrival remain unobserved. This means that information on short durations should be taken from durations covering January 1. The information is biased if such spells of immigration differ from spells shorter than 1 year that do not include January 1.

The GBA/CRV files have been linked by a unique identifier to observations in the Regional Income Panel 1995–2000 (*Regionaal Inkomens Onderzoek*, RIO), created by CBS. RIO is a panel of 2 million households, containing some 5 million individuals, about one third of the population. Individuals leaving or (re-)entering the household leave or (re-)enter the panel. The original GBA/CRV file includes about 600,000 immigrants, from which about a third can be retrieved in the RIO panel, thus generating a GBA/CRV/RIO file of some 200,000 immigrants. Naturalised immigrants are maintained in the RIO sample. The resulting GBA/CRV/RIO dataset is called the *Immigrant panel*. Essentially, it covers a third of all immigrants who registered in the GBA between 1990 and 2000 provided they have not left before January 1, 1998, and it records socio-economic data for the years 1995–2000. The Immigrant panel includes 53,000 refugees and gives panel information on labour income, the number of weeks worked and on socio-economic classification. The classification is based on the dominant income source during the year: employee, self-employed, on disability, social assistance or unemployment benefit, other (mostly non-participating, without an individual income). Labour income itself is taken from fiscal records and has very high reliability.

The immigrant panel does not include information on level of education of immigrants. Therefore, we use two other sources. The first is a file for asylum migrants who applied for asylum between 1995 and 2000 (called IND/ITS file). The application document has an entry for all immigrants, but immigration officers consider education mostly irrelevant for their purpose and often do not bother to report it. Institute for Applied Social Sciences (ITS) Nijmegen has coded the information but only for refugees, for the period 1995–2000.[1] The second source is a register of the government employment agency Centre for Work. and Income (CIW) that has assessed education for individuals who have contacted CIW to find a (new) job, obviously a very selective group.

The Immigrant panel contains about 53,000 refugees, of whom 43,000 satisfy our age constraint (15–59). We searched this sub-sample of 43,000 refugees in the IND/ITS and CWI file, using a unique common identifier and successfully found 16 339 refugees in both files. Some 27,000 refugees could not be traced in the IND/ITS file. The Immigrant panel contains those who registered in the GBA between 1990 and 2000, whereas the IND/ITS file contains asylum applicants from 1995 to 2000; checking the effect of this difference in time frame, we estimate that matching would be impossible for about 65% of the 27,000. The remaining part of the loss can be explained by the lag between the time of asylum application and the registration in the GBA, which is described below and no doubt some noise in the records. Finally, we matched the sample of 16,339 refugees with the CWI file. These matches provided two education variables: one from the IND/ITS file and another from the CWI file but both with substantial error. Below, we will assess the quality of information on education in detail.

2.2 Refugees

Asylum migrants (*refugees*) enter as applicants for asylum. Registered asylum migrants are immigrants who have been admitted and immigrants waiting for a decision on their asylum application. Admitted refugees are those who have obtained a title of residence; valid titles are temporary status (permission to stay until the situation in the home land is safe), A status (recognised as refugee and granted permanent residential status), "AMA" (admitted independents under 18) and admission for humanitarian reason. Admitted asylum migrants in principle are always registered in the GBA. Registration for asylum applicants is variable. If they are registered in the GBA at all, registration takes place several months after application. Since 1998, there are two special arrangements for asylum applicants. Under *Zelfzorgarrangementen* (Independent Housing), refugees find their own housing, with friends, relatives or otherwise. In this case, they will always directly be registered in the GBA. Under Central Housing, *Centrale Opvang voor Asielzoekers* takes care of housing. Asylum applicants in Central Housing are registered in the GBA when they obtain asylum status or after spending 1 year in Central Housing (since June 2000, after spending 6 months). Most applicants were registered when they left Central Housing. This means that the group of asylum migrants contains an unknown share of asylum applicants, i.e. an unknown mixture of admitted migrants and applicants for admission.

---

[1] We are very grateful to ITS for their generous offer to add their coding to our dataset.

2.3 Timing of events

The population register shows date of entry in the municipal register of population and keeps track of the immigrant's address. For the timing of events, we create three variables: years since migration (YSM), time spent waiting for a decision on the asylum application (*Statuswait*) and time spent in The Netherlands as an undocumented immigrant (*Undocyears*). The standard procedure for an asylum seeker is to be registered at the border by the Immigration Service, wait for a decision on the application for admission and in case of a positive decision, be registered in the population register as a resident. We denote the moment of registration by the Immigration Service as year of arrival, the moment of registration in the population register as year of settlement. We take registration in the population register as the moment of entry into Dutch society, as then a status has been granted and only then the immigrant is allowed to work and start building up rights to social security benefits. We measure YSM as time elapsed since settlement. The time elapsed between arrival and settlement is spent waiting for a decision on the application and is defined as Statuswait. In exceptional cases (3.6% in our sample), immigrants have been registered as settlers in the population register before they applied for admission as refugee with the Immigration Service. This could happen because of initial tolerance of undocumented immigrants: When the rules were tightened, these immigrants decided to apply for a formal status. We denote the time elapsed between registration in the population register (settlement) and status application (arrival) as undocumented years in cases where the former came before the latter. We should point out, however, as noted above, that the moment of registration in the population register was not always unequivocally defined. In perhaps 10% of the cases, applicants were registered in the population register while the decision on their application had not yet been taken. The definitions of these variables are summarised in Appendix B.

2.4 Selection of the sample

We will analyse data for refugees who are still present in The Netherlands in 2000, 13,436 out of 16,339; of these, we have 31,323 observations on the period 1995–2000. Following an entry cohort and, hence, using information on returned immigrants as well is not an attractive alternative, as it would only be feasible for cohorts entering in 1998 or later: It would restrict the analysis to fairly recently arrived immigrants only. We might also have opted for using all observations in the database up to their last moment of observation; final observations for individuals would then refer to 2000 or to year of departure if earlier. Our choice implies that we do not observe individuals that have left before 2000. This would be disturbing if return migration is selective. We are fairly confident that this is not the case, however. Our sample is restricted to those who have a valid permit to stay. We can observe departures for arrivals in 1998 or later. Among those with a permanent residential status in that sample, we only observe five people who have left (out of perhaps some 10,000 admissions). Those with a temporary permission to stay may be expelled when their homeland is declared safe (e.g. former Yugoslavia). In that case, return migration is an exogenous event and need not worry us.

By year of arrival, the sample spans the decade of the 1990s, but most observations date from 1996 or later: 3.6% arrived in 1990–1995, the remainder in 1996–2000. There is also some attrition from using the household as a sampling unit. The initial recording covers all members of a household; if someone later leaves the household, this means leaving the sample. Sample characteristics are given in Appendix A.

To create a reasonably homogenous sample, we require individuals to have a valid permit to stay, and we exclude individuals whose application is still being processed. As noted above, the sample also contains individuals who are still in the application process but are already registered at the GBA. This number is unknown but very small. The records contain many statements on the applicant's formal status, but there is no track record of progress in the decision-making process. Dates of decisions are not registered. Therefore, we decided to stay on the safe side and distinguish only three categories: *A status* (permanent permission to stay; includes also immigrants granted Dutch citizenship in 2001), *AMA* (entered as independent minor, i.e. not older than 18) and *preliminary status* (all other). Presumably, AMA refers to status upon entry, A status and preliminary status refer to the situation in 2000 (as last recorded status); status updates (by IND) occur, but the date of last recording is not known. Table 1 gives the distribution by status and country of origin.

In our sample of refugees, Iraq, Afghanistan and other countries each contribute about one fifth, 11% are from former Yugoslavia; Iran, Somalia, Sudan and the Soviet Union each contribute some 5–6 and 3% are from China. About two thirds of the refugees have a preliminary status, just more than a quarter has A status and 6% are AMA. AMAs are mostly from China and Somalia. Among the refugees with A status, Iran, Iraq and Afghanistan are over-represented.

3 Measuring education

We are specifically interested in the relevance of homeland education for socio-economic position after immigration, but we have reason to be suspicious about the reliability of recorded education. The original documents may register the applicant's education, but if so, registration is not according to a standardised classification system. ITS analysts have coded the entries to a standard classification in nine

Table 1 Admissions by title of residential status and country of origin

	A status (%)	AMA	Preliminary	Total ($N=100$)
Iran	42.14	0.74	57.12	674
Iraq	35.51	0.32	64.17	3,123
Somalia	15.90	15.33	68.76	874
China	1.17	53.38	45.45	429
Afghanistan	37.19	1.05	61.76	2,947
Sudan	28.45	4.38	67.17	594
Former Yugoslavia	27.52	0.16	72.33	1,272
Soviet Union	28.08	2.91	69.01	755
Other countries	16.11	13.51	70.38	2,768
Total	28.36	6.20	65.44	13,436

categories (see Appendix B). From analysis by ITS, we know that education is missing in many cases and that there is reason for doubting the reliability of the recorded levels. We also know that education is not an important variable in the decision process and that immigration officers have no special interest in it. In fact, they consider it irrelevant and often ignore it. Hence, before attempting any analysis, we should assess the quality of measurement.

Table 2 presents the distribution by education levels, distinguished by country of origin. The first thing to note is that in 45% of the cases, education is missing. Seven percent has no education at all, 23% has basic (including extended basic), 14% has secondary and 11% has tertiary education. If among missing recordings individuals have the same distribution by education as those who are observed (which may well be true, see below), 13% of all refugees would have no education at all, and more than half (55%) would have no more than extended basic education. Fifteen percent would have a higher education. This points to a rather unequal distribution of education.

Refugees from China, of whom many are AMA, have remarkably low levels of education and so have refugees from Somalia. Among the refugees from Iran, there is a remarkably high share with secondary education; Sudan has relatively many highly educated refugees, and the distribution from the Soviet Union is rather bimodal: high shares with extended lower and with high education. Refugees from Iraq are often well educated. In the total sample, 20% has primary education, 27% secondary, 8% tertiary and for 45% education is missing. By title of residential status (not shown here), refugees with A status have higher average education level, and AMAs have lower average level of education. Among all refugees, 28% has A status, whereas among refugees with tertiary education, 43% has A status.

We also have observations on education recorded by the CWI, the public employment service that assists individuals in finding a job. Registration as job seeker is a requirement for obtaining social benefit. Clearly, this registration is highly selective. However, we might assume that employment service agents are more dedicated in registering education, as it is an important instrument for the service they have to provide: They have an interest in accurate assessment. However, they might also apply censoring and only register education they consider relevant for the Dutch labour market. We do not know whether individuals have obtained any additional education in The Netherlands. Upon a first visit to the employment agency, this seems unlikely, but with later visits, an update might have taken place.

From Table 3, we may note first of all that the missing observations do not match: They are not concentrated as single diagonal entry in the cell (missing IND and missing CWI). Missing observations must result from different processes in the two agencies and are not a unique property of the respondent. The overall proportions are about equal, at 45% for IND and 49% for CWI, but this must be coincidence, as IND missings are due to non-registration by the immigration officer, whereas CWI missings must be due to absence of contact with the employment service. Interestingly, the proportion of missing observations on IND education is virtually the same for every level of CWI education. If we are justified in assuming that CWI registration is reasonably reliable, this would imply that missing observations in IND are unrelated to the level of education and hence that the distribution of observed education is representative for all refugees: We can relate the frequencies to only those individuals for whom education has been registered.

Table 2 Education level by country of origin, percentages (IND records)

	None	1–3 basic	4–5 basic	Basic	Extended basic	Secondary, general	Secondary, vocational	Some tertiary	Tertiary	Missing	Total
Iran	2.82	0.89	1.34	2.23	10.09	23.74	2.23	3.12	10.98	42.58	674
Iraq	4.67	1.44	2.34	4.13	9.54	8.36	3.30	3.39	12.33	50.50	3,123
Somalia	20.82	8.92	5.38	7.32	11.90	10.64	0.80	0.57	2.52	31.12	874
China	4.20	15.38	20.05	14.22	13.99	4.66	0.93	0.00	0.23	26.34	429
Afghanistan	5.09	1.26	2.21	4.21	4.48	9.54	0.71	3.02	9.74	59.76	2,947
Sudan	5.72	1.52	2.86	5.05	8.92	11.11	2.02	5.56	18.52	38.72	594
F. Yugoslavia	6.53	0.94	2.99	4.64	19.03	12.03	10.46	2.04	2.67	38.68	1,272
F. Soviet Union	4.90	2.25	1.85	4.77	16.16	16.16	5.56	1.85	13.51	32.98	755
Other countries	9.68	4.34	6.11	10.91	12.17	11.92	1.88	2.13	4.34	36.52	2,768
Total (%=100)	6.97	2.90	3.86	6.10	10.54	11.06	2.90	2.63	8.45	44.60	
Total (N)	937	390	518	820	1,416	1,486	389	353	1,135	5,992	13,436

Table 3 Education IND and education CWI

| Education IND | Education CWI | | | | | | | |
	Unknown	Basic	Lower (vocational+ general)	Intermediate (vocational+ general)	Higher vocational	University	Missing	Total
No education	9.09	9.99	5.69	1.33	1.00	0.66	8.68	6.97
1–3 year primary	3.74	4.17	2.97	0.96	0.14	0.26	3.41	2.90
4–5 year primary	5.88	5.78	4.17	1.33	0.14	0.40	4.35	3.86
Primary	4.81	8.47	7.08	3.91	0.29	0.26	6.80	6.10
Extended primary	13.9	11.68	14.74	9.14	1.85	1.32	11.3	10.54
Secondary, general	6.95	9.49	13.35	20.58	9.56	4.50	10.13	11.06
Secondary, vocational	0.53	2.57	3.67	5.68	3.71	1.06	2.44	2.90
Some tertiary	2.67	1.48	1.64	5.01	5.71	5.17	2.16	2.63
Tertiary	8.56	3.20	3.23	8.85	28.96	40.26	5.63	8.45
Missing	43.85	43.17	43.45	43.22	48.64	46.09	45.10	44.6
Total (N=100)	187	2,372	1,581	1,356	701	755	6,484	13,436

There is certainly no agreement between IND and CWI on individuals' level of education. Table 4, with education registered in three comparable levels, shows this quite clearly. If we consider only cases for which both institutions record an education level (i.e. exclude missing observations), the diagonal elements in Table 5 would be 0.50, 0.65 and 0.65, meaning that for the given classification by CWI, in no more that two thirds of the cases would IND record the same level.

With levels of education grouped in primary, secondary and tertiary (to allow for matching classifications), we can calculate that in 6.8% of all cases, the IND level is higher than the CWI level, whereas in 5.1%, the reverse holds (13,436 cases, 2,593 with primary, 2,966 with secondary and 1,467 with tertiary education, CWI classification). This points to some upward bias in the IND registration relative to the CWI registration, as one might have anticipated: IND is the individual's assessment without any check; CWI coding is based on the registration by an

Table 4 Education IND and education CWI, three levels

| IND | CWI | | | | |
	Primary	Secondary	Tertiary	Missing	Total
Primary	28.06	14.20	1.58	23.24	19.83
Secondary	25.13	36.64	16.28	26.03	27.12
Tertiary	3.60	5.82	34.82	5.63	8.45
Missing	43.22	43.34	47.32	45.10	44.60
Total	100	100	100	100	100

Table 5 Determinants of employment and benefits: coefficients and relative risk ratios (RRR) of the multinomial logit model; estimated using the pooled sample

	Employment			Benefits		
	Coefficient	Std. Error	RRR	Coefficient	Std. Error	RRR
Age	−0.03***	0.003	0.97	0.05***	0.003	1.05
Woman	−1.29***	0.054	0.28	0.01	0.051	1.01
YSM	1.27***	0.094	3.54	1.24***	0.091	3.46
YSM2	−0.12***	0.028	0.89	−0.25***	0.023	0.78
YSM3	0.01**	0.004	1.01	0.02***	0.003	1.02
YsmIraq	−0.09	0.075	0.91	0.00	0.077	1.00
YsmSomali	−0.28***	0.085	0.75	0.06	0.091	1.06
YsmChin	0.54***	0.145	1.71	0.34***	0.112	1.41
YsmAfgh	0.15*	0.079	1.16	0.52***	0.082	1.69
YsmSudan	0.16	0.110	1.18	0.32***	0.123	1.38
YsmYugos	0.04	0.080	1.04	−0.05	0.094	0.96
YsmSovU	−0.07	0.096	0.93	−0.23**	0.093	0.79
YsmOther	−0.09	0.071	0.92	−0.18**	0.079	0.83
arrival95 (reference)						
arrival96	0.42***	0.084	1.52	0.44***	0.083	1.56
arrival97	0.65***	0.094	1.92	0.18**	0.093	1.20
arrival98	0.64***	0.101	1.89	−0.56***	0.104	0.57
arrival99	0.15	0.145	1.16	−0.87***	0.149	0.42
arrival00	0.44	0.316	1.55	−0.49	0.333	0.61
No education (reference)						
edu1_3 y	0.11	0.171	1.12	0.20	0.184	1.22
edu4_5 y	0.20	0.155	1.23	0.22	0.148	1.24
EduPrim	0.38***	0.135	1.46	0.15	0.137	1.16
EduPrim_ext	0.53***	0.119	1.70	0.49***	0.117	1.64
eduSec_gen	0.45***	0.120	1.56	0.48***	0.114	1.62
eduSec_voc	0.79***	0.165	2.21	0.44***	0.156	1.55
EduHigh_some	0.76***	0.169	2.13	0.39**	0.169	1.48
EduHigh	0.67***	0.128	1.94	0.47***	0.119	1.60
EduMiss	0.38***	0.104	1.47	0.14	0.100	1.15
Iraq	1.06***	0.197	2.89	0.16	0.186	1.18
Somalia	1.13***	0.223	3.10	−0.14	0.232	0.87
China	−1.20***	0.382	0.30	−0.39	0.268	0.68
Afghan	1.09***	0.196	2.97	−0.45**	0.189	0.64
Sudan	0.99***	0.231	2.68	0.10	0.234	1.11
Yugoslavia	0.76***	0.213	2.14	−0.07	0.216	0.94
SovietUni	0.32	0.240	1.38	1.10***	0.207	3.00
Other	0.20	0.200	1.22	0.22	0.191	1.24
Iran (reference)						
A_Status	1.02***	0.066	2.78	2.47***	0.062	11.85
AMA	0.03	0.099	1.03	1.86***	0.114	6.43
Undocyears	0.56***	0.035	1.75	0.15***	0.036	1.17
Statuswait	−0.68***	0.093	0.51	−0.31***	0.077	0.73
Naturalised	0.82***	0.075	2.27	0.81***	0.075	2.25
Returned	−1.11***	0.189	0.33	−1.39***	0.278	0.25
Married	−0.14**	0.056	0.87	0.54***	0.053	1.72
Amsterdam	0.16	0.096	1.17	0.68***	0.100	1.97
Rotterdam	−0.44***	0.113	0.65	0.10	0.108	1.10
The Hague	−0.26**	0.110	0.77	0.22**	0.102	1.24
Utrecht	0.18	0.159	1.20	0.07	0.200	1.07
Constant	−3.45***	0.252	0.03	−5.03***	0.242	0.01
Number of persons	13,436					
Number of person-years	31,323					
Log Likelihood	−22,103					
$\chi^2(92)$	7,588					

The multinomial logit model estimated has three outcomes: employment, benefits and non-participation, which is the reference category. Standard errors are adjusted for clustering on person ID.

*$p<0.1$; **$p<0.05$; ***$p<0.01$ (Standard errors are robust)

employment agency that has interest in accurate assessment. The employment agents translate foreign education into the guessed Dutch equivalent and might perhaps be inclined to some downward bias because of unfamiliarity with foreign schooling systems. However, the bias is quite modest, which lends credibility to the IND data.

We have analysed possible patterns of non-recording of education by IND officers by running a logistic regression. Registration of education is indeed related to some variables: Education is more often registered for immigrants who are older at arrival and for men, it is better known for later arrivals, and there are significant differences between countries of origin: better known for China, Soviet Union and Somalia, less often known for Afghanistan and Yugoslavia. "Undocyears" (years spent without residential permit) and "Statuswait" (time spent waiting for a decision on the application) also have significant, negative effects on the probability of registration.

We have made inquiries with IND and with immigration officers who do the intake interviews and registration of immigrants. They could not give any explanation on the pattern of registration of education, and they are absolutely unaware of any systematic effects.

We conclude that non-recording of level of education by IND has different incidence by country of origin, years of arrival and gender of the refugees, but there is no indication of a systematic rule applied by immigration officers. We do not see any reason to fear that non-registration of education is related to level of education. Neither do we see any indication that non-registration or erroneous registration would be related to unobserved ability of applicants. From comparing IND and CWI registration, we conclude that there is evidence of only modest upward bias in the level of education recorded by IND. However, the substantial variation in the cross-classification of the two registrations indicates that measurement error in the level of education is far from negligible.

We have considered using information on homeland occupation (also coded by ITS) as a variable to assess the reliability of registered education. However, a cross-tabulation of education and occupation shows wide dispersion of education by occupation. Moreover, many educations are so low and so little specific that it would be hard to use the additional information to test the reliability of education. There are auto mechanics and farm hands with tertiary education and pharmacists with just extended basic education. The matrix is simply too far removed from diagonality to yield useful additional information. A logistic regression shows no relation between recording education and recording occupation.

With more than one measurement of an individual's education, it is possible to deduce information on the magnitude and effect of measurement errors. Kane et al. (1999) found that problems are more severe for incomplete educations than for completed degrees. Just as Battistin and Sianesi (2006), they stressed that with categorical data, measurement error cannot be classical, as the upper and lower limits on education imply that errors depend on true level of education. Below, we deduce some information on the possible magnitude of measurement error in the continuous case, with education measured in years. In our estimates in Sections 4 and 5, we test robustness of results by restricting the sample to cases where the two measures of education are identical. This, of course, does not fully exploit all the information in the data, but neither does it require additional assumptions (Kane et al. assumed independent measurement errors, which may well be violated in our case).

Formalising, we have two measures of an individual's education, S_{iT} as measured by IND and S_{iC} as measured by CWI, both measured in years (we have translated education levels to years, as specified in Appendix B). Assume:

$$S_{iT} = S_i + e_{iT} \tag{1}$$

$$S_{iC} = S_i + e_{iC} \tag{2}$$

Both measurements report individual i's true education but with different measurement errors. Assuming that the errors are independent of the true value, we can write for the variances:

$$V_T = V + V_{eT} \tag{3}$$

$$V_C = V + V_{eC} \tag{4}$$

where V measures the variance of true education across individuals, and V_e measures the variances in the error terms.

We can write the covariance V_{CT} as

$$\begin{aligned} V_{CT} &= E\{(S_{iT} - E(S_{iT}))(S_{iC} - E(S_{iC}))\} \\ &= E\{(S_{iT} - E(S_i))(S_{iC} - E(S_i))\}, \end{aligned} \tag{5}$$

under the assumption of zero-expected measurement error and independence of true education levels. Substituting the definitions (1) and (2), we can write this as

$$V_{CT} = V + \rho\sqrt{V_{eT}V_{eC}} \tag{6}$$

where ρ is the correlation between e_{iT} and e_{iC}. From the three Eqs. 3, 4 and 6, we can identify the three variances if we know (or make assumptions on) the correlation between the two measurement errors. The variance in IND education is 16.50, the variance in CWI education is 13.67 and the covariance is 8.38. With these numbers, substituting Eqs. 3, 4 and 6 and squaring, we can solve the resulting quadric equation for V. Solving the equations for given values of the correlation coefficient gives the results plotted in Fig. 1. Measurement errors in IND recordings are always larger than in CWI recordings, as one would anticipate. As the squaring may permit solutions that are not solutions to Eq. 6 itself, we check whether solutions to the quadratic also obey Eq. 6 itself.[2] It turns out that the positive root holds for negative correlation and the negative root holds for positive correlation. It seems fair to rule out negative correlation between measurement errors, as it would be hard to explain (an employment officer "punishing" an immigrant for lying to the IND?). Intuitively, it seems reasonable to assume that the correlation will be somewhere in the range 0 to −0.7. Then, with correlation 0, true variance would be 8.38, IND measurement variance 8.12 and CWI measurement variance 5.29. With correlation at 0.7, V would be 11.03, V_{eT} 5.46 and V_{eC} 2.63. The values imply that in the IND records, the ratio $V/(V+V_{eT})$ runs between 0.51 and 0.67. In an ordinary least squares (OLS) regression with years of education as single explanatory variable, this ratio would give the estimated regression coefficient as a proportion of the true coefficient. With

---

[2] We are grateful to the referee who corrected the neglect of this condition in our original version.
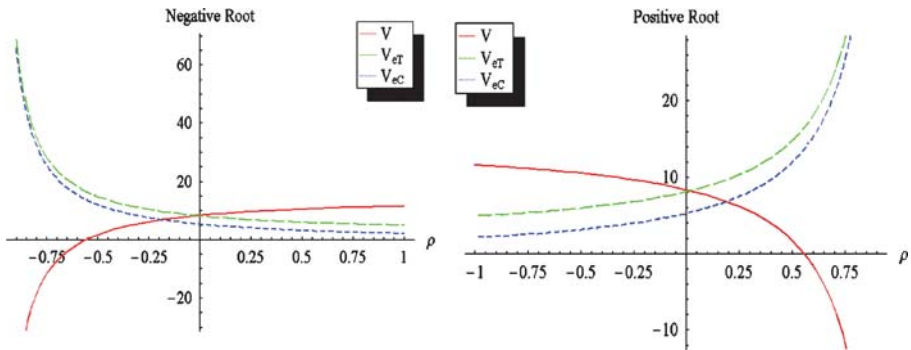
Fig. 1 Variances as a function of the correlation coefficient

more than one explanatory variable, all coefficients are biased downwards, but it is hard to determine the magnitudes (Wooldridge 2002, p. 75). Unfortunately, in logit and probit regression, the effect of measurement errors on estimated coefficients is unknown, and one cannot assume analogy to linear regression.[3] The indication of substantial measurement errors in education is reason for concern. We will use the double measurement of education to check the reliability of our results.

4 Socio-economic status

We have estimated the effect of education and other variables on socio-economic status, distinguishing three states: $Y_i=1$ if individual i is employed as employee or self-employed, $Y_i=2$ if person i does not work and receives some social benefit (unemployment, welfare) or $Y_i=3$ if individual i neither works nor receives a benefit (non-participating). Considering these three states, we have estimated a multinomial logit model using pooled data, using all available observations for a given individual, in different years (with correction of the standard errors for repeated observations per individual), as a panel estimation of a multinomial logit model is hard to construct and estimate (we have also estimated two sets of random effect panel logit models, with work versus non-work, benefits versus non-benefits, and work versus non-participation, benefits versus non-participation; the key conclusions are similar). Non-participation is defined as the reference category. Unemployment benefit and welfare are not distinguished, as the number of individuals receiving unemployment benefit is very small (only 56 person-years) during the whole panel period 1995–2000. The model has been estimated on all individuals who are present in 2000. We do not include refugees who have returned home before 2000 or apply correction for such attrition. As noted above, among those admitted permanently (with an A status), no one leaves, and among those admitted temporarily, departures are exogenous, dictated by the political situation in their homeland.

Admitted immigrants have identical entitlements to social security benefits as the native Dutch. However, unemployment benefits are conditional on work history,

---

[3] Private communication, professor J.S. Cramer, University of Amsterdam. As Cramer (2003, Section 5.3) shows, omitted variables lead to a downward bias in discrete models.

which will disqualify immigrants in the early years after arrival. Social assistance does not depend on length of stay in The Netherlands but is means-tested at the household level and may disqualify marital partners or children (although the level of the benefit will depend on household composition). Refugees are provided shelter, food and a small amount of cash while their application is in process.

Formally, individual i's contribution to the likelihood function in state j is

$$P(Y_{it} = j | \mathbf{X}_{it}) = \frac{e^{\mathbf{X}_{it}\beta_j}}{1 + \sum_{k=1}^{2} e^{\mathbf{X}_{it}\beta_k}} \tag{7}$$

where $\mathbf{X}_{it}$ is a vector of explanatory variables for individual i in year t, β_j is a vector of coefficients, varying with three alternative outcomes and the coefficients for reference outcome 3 (non-participation) are normalised to zero. The time subscript on $\mathbf{X}$ is actually excessively general: Only YSM and city of residence are time varying; all other variables are measured at arrival. We estimate probabilities of work and receiving benefits using the pooled sample of the panel period 1995–2000. Because each individual contributes more than once to the sample, standard errors are adjusted for the intra-individual correlation (we give robust standard errors). We do not correct standard errors for possible intra-household correlations because asylum migrants are usually single (male) individuals. Because the estimated coefficients of a multinomial logit model are difficult to interpret directly, we report relative risk ratios as well. The relative risk ratio measures the effect of a variable on the probability of outcome j relative to the probability of the reference outcome.

In Table 5, we report estimation results for the probability of work in the most extended specification. We control for cohort effects with dummies for year of arrival and note that although we also include the time profile of YSM, earlier arrivals do better. Differences between countries of origin are marked. For some countries, there are differences in level only; for other countries, the interactions between country dummy and YSM are also significant. We will return to country effects in Section 6. Among the residential locations, Rotterdam and The Hague stand out with lower probabilities of employment. Age at arrival has a negative effect (age squared, when added, had insignificant effect), women are less likely to work than men. We have also considered the effect of the situation in the year after our observation interval. Those who will be naturalised in 2001 are more likely to work, and those who will then have returned (or administratively removed) are less likely to work. The latter result hints at selectivity in return migration, in spite of our earlier remarks. Undoc years have a positive effect: Refugees who have been in The Netherlands as undocumented workers before reporting to IND have a higher probability to work. This is as anticipated: As undocumented workers, they will mostly have worked, and effectively, this adds experience to their YSM. Statuswait has a negative effect. Spending more time in the application procedure reduces the probability to work, even after controlling for the other duration variables.

YSM, i.e. time elapsed since registration in the population register GBA, has a monotonic positive effect on the probability to work, as one might expect. We estimated a cubic function for YSM to obtain maximum flexibility, although this cannot be extrapolated very far, as we only estimate over a 5-year interval. We have plotted the profile in Fig. 2.
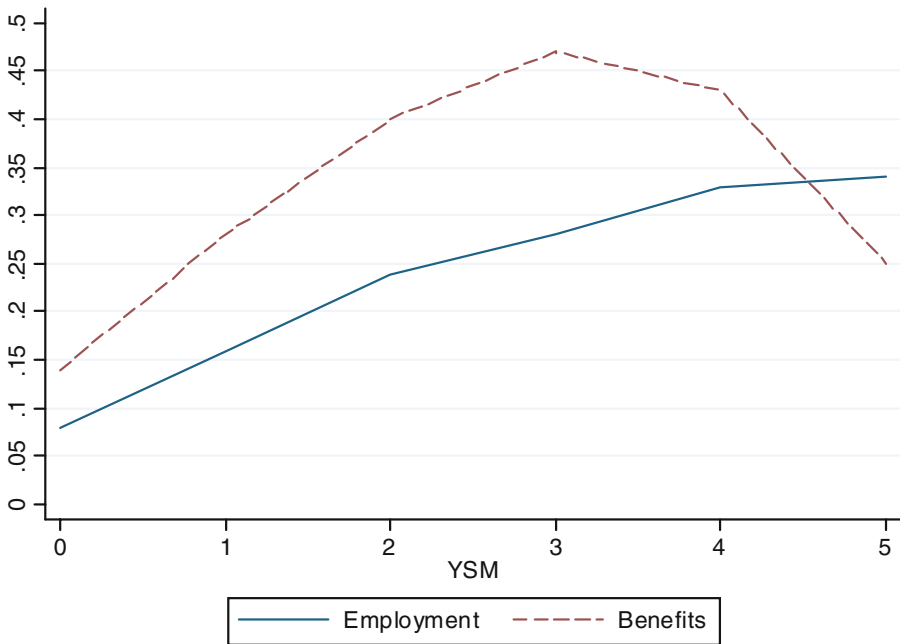
Fig. 2 Predicted probability of employment and benefit status over time elapsed in The Netherlands (years since migration YSM)

Our special interest is in the effect of home country schooling. Generally, there is a positive effect of schooling on the probability to work. Considering magnitudes and significance levels, one may distinguish three steps: less than basic education, basic up to secondary general and secondary vocational and higher. The *probability of work* increases markedly between steps and is quite similar within the steps. Interestingly and perhaps not surprisingly, the effect of secondary level schooling is split: Secondary vocational education has a stronger effect on the *probability of work* than secondary general education. This suggests that vocational skills are directly transferable, whereas general academic skills are not. Within the highest step, the effect of education diminishes slightly, although not significantly. Before checking the reliability of the results for education, we will discuss the results for benefit recipient status.

The probability of receiving a social benefit differs by country of origin, increases for more recent arrival cohorts, is higher for older and married individuals but not different for women and, among cities, is highest in Amsterdam. Undocyears has a positive effect, whereas Statuswait has a negative effect. Refugees who are no longer present in 2001 have a lower probability of recipient status and refugees who have become naturalised in 2001 have a higher probability; these results also indicate that individuals who are less integrated into Dutch society are more likely to leave.[4]

---

[4] The status of naturalised and returned may have materialised subsequent to measurement of the dependent variable. In this case, including these variables may bias coefficient estimates. Nevertheless, we include these informative variables because the real decision on naturalisation and returning home may well have been made earlier. Besides, the impact of omitting these variables on coefficient estimates has been proved to be negligible.

YSM has a positive effect that peaks at 3 years (Fig. 2). The effect is initially positive, as refugees have to build up entitlements (for example, if they have been registered in GBA before obtaining residential status, they will only be entitled to social assistance once they have obtained their status). Later, the profile declines as the immigrants find jobs.

The effect of education is essentially a single jump at extended primary education. Individuals with education below that level have the same probability of receiving benefits; individuals with that level or higher share a significantly higher benefit probability. One might perhaps have anticipated that the effect would be negative, as those with a higher education would be better able to find a job. It is not quite clear how to interpret this unanticipated finding. It may be a reflection of the build-up of benefit entitlement with work history, as the results for work and benefit are to some extent parallel: Higher education leads to higher probability of work, which also leads to increases in the probability to be eligible for social benefits. However, there are very few individuals who receive unemployment benefits. Perhaps the means testing for entitlement of assistance is less binding for the higher educated. The highly educated individuals may be better at negotiating to get access to the social system.

As we are specifically interested in the effects of homeland education, we have concentrated our analysis of reliability on the effects of education. We have made separate estimations on sub-samples, and we attempted to allow for the reliability of recorded levels of education. In Table 6, we present estimates of probabilities for employment and benefit, on five sub-samples: arrivals 1995, arrivals 1998–2000, men, women and excluded if education missing. A generally positive effect of level of education on the probability of employment but without university education ranking on top is confirmed in all these regressions. The estimates confirm a strong positive effect of secondary vocational education. The parabolic effect of education is more noticeable in the 1995 cohort (model II) than in the 1998–2000 cohort (model III). This is rather surprising, as it suggests that lagging employment probability of the university educated increases over time; one might have anticipated the reverse effect. The results for the probability to receive social benefits estimated on sub-samples also basically repeat the findings for the full sample: There is no monotonic relationship between benefit recipient status and level of education. The relatively stronger effects for women compared to those for men (compare model V with model IV) might have some relationship with building up unemployment benefit entitlement from work (the effect of education on work is also stronger), but it may also be that more highly educated women are less sensitive to the means test restriction than men are.

To test the sensitivity of our results for the reliability of recorded education, we have made selections based on the combination of IND and CWI records (Table 7), by discarding "unreliable" recordings of education. We defined "unreliable" as clear mismatch in the two measurements. The first selection rule we applied, reported under model II, is the following:

IND primary or less   Accepted if CWI classification Basic
IND extended primary and secondary   Accepted if CWI lower or intermediate

Table 6 Testing on sub-samples: selected coefficients of multinomial logit models

Variable	Model I	Model II	Model III	Model IV	Model V	Model VI
Employment						
edu1_3 y	0.11	0.07	−0.40	0.08	0.06	0.12
edu4_5 y	0.20	0.14	0.13	0.03	0.42	0.21
EduPrim	0.38***	0.50*	−0.14	0.40**	0.17	0.39***
EduPrim_ext	0.53***	0.63**	0.32	0.41***	0.68***	0.54***
eduSec_gen	0.45***	0.75***	0.51**	0.31**	0.72***	0.46***
eduSec_voc	0.79***	1.37***	0.41	0.61***	1.18***	0.72***
EduHigh_some	0.76***	0.87**	0.48	0.48**	1.28***	0.74***
EduHigh	0.67***	0.83***	0.69***	0.50***	0.97***	0.63***
EduMiss	0.38***	0.44**	0.35*	0.26**	0.48**	
Benefits						
edu1_3 y	0.20	−0.32	0.44	0.11	0.27	0.21
edu4_5 y	0.22	0.44	0.08	−0.02	0.46**	0.23
EduPrim	0.15	0.34	−0.10	0.31	−0.10	0.14
EduPrim_ext	0.49***	0.83***	0.22	0.40**	0.57***	0.50***
eduSec_gen	0.48***	0.85***	0.11	0.41**	0.54***	0.47***
eduSec_voc	0.44***	0.94***	−0.64	0.39*	0.39	0.46***
EduHigh_some	0.39**	0.53	0.07	0.08	0.74***	0.39**
EduHigh	0.47***	0.70***	0.08	0.38**	0.54***	0.48***
EduMiss	0.14	0.26	0.10	0.05	0.20	
Number of persons	13,436	1,645	5,599	8,453	4,983	7,444
Number of person-years	31,323	5,677	8,612	19,651	11,672	17,471

The multinomial logit models are estimated with the same variables as in Table 5 in addition to the education variables presented in Table 6, but their coefficients are not presented here. Non-participation is the reference category. Robust (clustered) standard errors.

Model definitions: *Model I*, reference (full sample); *Model II*, sub-sample of individuals arrived in 1995 according to IND; *Model III*, sub-sample of individuals arrived between 1998 and 2000 according to IND; *Model IV*, sub-sample of men; *Model V*, sub-sample of women; *Model VI*, sub-sample of individuals for whom education is known (non-missing)

*$p<0.1$; **$p<0.05$; ***$p<0.01$

IND secondary vocational, some tertiary   Accepted if CWI intermediate or higher vocational

IND higher   Accepted if CWI university

Because of the difference in the two classification systems, this is not a strict criterion to a perfect match, so some noise is inevitably left. We therefore also used as an alternative selection rule that the classifications should agree on the level of primary/secondary/tertiary. We then estimated two specifications: the usual specification with all IND categories (model III) and a specification with three levels only (primary/secondary/tertiary; model IV). By requiring a credible match between IND and CWI classification, we reduce the sample to those observations for which CWI classification is available. This is quite restrictive and certainly not random. Therefore, we also re-estimated the non-restricted versions on the sub-sample for which both IND and CWI education levels are available. This is reported as model I.

The restriction of the sample to individuals with observations from IND and from CWI has an effect on estimated coefficients: Significant coefficients increase for the employment equation (often by 30% or more) and decline in some specifications for

Table 7 Restricting the sample to matching education: selected coefficients of multinomial logit models

Variable	Model I	Model II	Model III	Model IV
Employment				
edu1_3 y	0.29	0.70**	0.70**	
edu4_5 y	0.33	0.58**	0.57**	
EduPrim	0.51***	0.78***	0.66**	
eduPrim_ext	0.68***	0.94***	0.95***	
eduSec_gen	0.66***	0.85***	0.93***	
eduSec_voc	0.83***	1.06***	1.62***	
eduHigh_some	0.97***	1.08***	1.53***	
EduHigh	0.83***	1.04***	1.12***	
eduINDsec				0.62***
eduINDter				0.65***
Benefits				
Edu1_3 y	0.29	0.52*	0.31	
Edu4_5 y	0.33*	0.36*	0.27	
eduPrim	0.14	0.25	0.22	
eduPrim_ext	0.50***	0.52***	0.35*	
eduSec_gen	0.52***	0.47***	0.46***	
eduSec_voc	0.42**	0.43**	0.60**	
eduHigh_some	0.46**	0.59**	0.42	
eduHigh	0.49***	0.50***	0.45**	
eduINDsec				0.27**
eduINDter				0.29*
Number of persons	3,884	2,993	2,301	2,301
Number of person-years	10,637	8,337	6,354	6,354

The models are estimated using the sub-samples described with the same variables as in Table 5 in addition to the education variables presented in Table 7, but their coefficients are not presented here. Non-participation is the reference category. Robust (clustered) standard errors

Sub-samples are defined according to the following rules: *Model I*, all IND categories if CWI classification is available; *Model II*, all IND categories; observations excluded according to the 'first selection rule' in the text: *Model III*, all IND categories if the IND and CWI classifications are agreed on three levels; *Model IV*, three levels IND categories if the IND and CWI classifications are agreed on three levels

*$p<0.1$; **$p<0.05$; ***$p<0.01$

the benefit equation. However, two key conclusions remain unaffected. The probability of employment increases in steps with the level of education but is not highest for the highest level of education. The probability of receiving benefit is not monotonically declining in the level of education but if anything is closer to a parabolic relationship. If we restrict education levels to (matching) primary, secondary and tertiary only, we find that employment and benefit probabilities are equal for secondary and tertiary educated individuals and above the probabilities for those with primary education only.

5 Earnings

In Table 8, we present estimates for earnings for employees; that is, individuals for whom labour earnings are the most important source of income during the year. It is the natural logarithm of annual labour income divided by weeks worked and deflated

Table 8 Panel GLS random effect estimations of log weekly earnings, 1995–2000

Weekly wages		Model I	Model II	Model III	Model IV	Model V	Model VI
Age		0.04***	0.06***	0.03***	0.04***	0.05***	0.04***
Woman		−0.46***	−0.43***	−0.30***			−0.43***
YSM		0.26***	0.37***	0.24**	0.26***	0.22**	0.21***
arrival96		0.16***			0.16***	0.09	0.13**
arrival97		0.18***			0.17***	0.13	0.11
arrival98		0.18***			0.16**	0.19	0.03
arrival99		−0.09			−0.11	−0.14	−0.20
arrival00		0.65**			−0.10	1.17**	0.63*
edu1_3 y		0.00	0.62***	0.15	−0.07	0.27	0.00
edu4_5 y		0.21*	0.45**	0.16	0.16	0.26	0.23**
EduPrim		0.28***	0.49***	0.11	0.31***	0.00	0.28***
eduPrim_ext		0.36***	0.74***	0.38*	0.32***	0.41**	0.39***
eduSec_gen		0.31***	0.51***	0.17	0.32***	0.10	0.34***
eduSec_voc		0.21*	0.45**	−0.33	0.18	0.31	0.21*
eduHigh_some		0.18	0.20	0.22	0.18	0.06	0.23*
EduHigh		0.20**	0.18	0.21	0.21**	−0.03	0.24***
EduMiss		0.13*	0.20	0.28	0.14	−0.01	
Iraq		0.35**	0.13	0.27	0.41**	0.08	0.08
Somalia		0.65***	0.61**	0.75**	0.66***	0.50	0.51**
China		−0.23	−0.56	0.53	−0.10	−0.79	−0.51
Afghan		0.29**	0.18	0.31	0.35**	0.03	0.14
Sudan		0.61***	1.40***	0.17	0.65***	0.75	0.53**
Yugoslavia		0.61***	0.89***	0.27	0.61***	0.55*	0.46**
SovietUni		0.38*	0.30	−0.19	0.26	0.44	0.35
OtherC		0.51***	0.88***	0.30	0.48***	0.63**	0.33
A_Status		0.02	−0.03	0.02	0.06	−0.10	0.00
YsmIraq		−0.05	−0.02	−0.05	−0.04	−0.04	0.06
YsmSomali		−0.08	−0.12	−0.13	−0.05	−0.15	−0.04
YsmChin		0.17*	0.20	−0.37	0.10	0.44**	0.27**
YsmAfgh		−0.03	−0.01	−0.09	−0.04	0.02	0.03
YsmSudan		−0.02	−0.28***	0.11	−0.01	−0.18	0.04
YsmYugos		−0.08*	−0.21***	−0.12	−0.07	−0.12	−0.02
YsmSovU		−0.03	0.02	0.15	0.03	−0.10	0.00
YsmOther		−0.05	−0.20**	−0.13	−0.03	−0.10	0.02
AMA		0.10	0.32**	−0.67**	0.12	−0.11	0.16
Undocyears		0.13***	0.15***	0.05	0.14***	0.08	0.12***
Statuswait		−0.14***	−0.16	−0.09	−0.19***	−0.04	−0.14**
Married		0.19***	0.19**	0.11	0.18***	0.19**	0.23***
Amsterdam		0.08	0.28**	−0.12	0.03	0.46***	0.02
Rotterdam		0.14*	0.39***	0.31	0.13	0.16	0.22**
DenHaag		0.16**	0.26	0.26	0.15*	0.29	0.08
Utrecht		0.03	0.03	−0.10	0.02	0.17	−0.02
Constant		2.36***	1.56***	3.03***	2.37***	2.10***	2.60***
σ_μ		0.70	0.69	0.37	0.70	0.66	0.70
σ_ε		0.71	0.61	1.06	0.71	0.71	0.67
η		0.49	0.56	0.11	0.50	0.46	0.52
R-sq	Within	0.10	0.23	0.01	0.10	0.10	0.12
	Between	0.24	0.40	0.13	0.20	0.32	0.25
	Overall	0.25	0.40	0.12	0.20	0.31	0.26
N-person		3,703	737	859	2,926	777	1,882
N-person-year		5,933	1,424	1,045	4,762	1,171	3,088

The dependent variable is the logarithm of weekly earnings for the period of 1995–2000 in the prices of 1995. σ_μ is standard deviation of individual random effect, μ_i; σ_ε is standard deviation of the idiosyncratic errors, ε_{it}; η is fraction of total variance because of ε_{it}.

Definitions of sub-samples: *Model I*, basic specification; *Model II*, arrivals 1995 (IND registration); *Model III*, arrivals 1998–2000 (IND registration); *Model IV*, men only; *Model V*, women only; *Model VI*, observations with education missing deleted

*p<0.1; **p<0.05; ***p<0.01

by cost-of-living (base year 1995). We estimate the following panel generalised least squares (GLS) model with random effects (Wooldridge 2002, ch.10).

$$y_{it} = \boldsymbol{X}_{it}\beta + \mu_i + \varepsilon_{it} \quad t = 1, 2, \ldots, T \tag{8}$$

Where y_{it} is earnings of individual i in period t, x_{it} is a vector of individual characteristics, β is the corresponding coefficient, μ_i is the individual random effect, which is assumed to be uncorrelated with x_{it} and to satisfy $E(\mu_i) = 0$, $E(\mu_i^2) = \sigma_\mu^2$ and ε_{it} is idiosyncratic errors that are strictly exogenous, $E(\varepsilon_{ti}) = 0$, serially uncorrelated, $E(\varepsilon_{it}\varepsilon_{is}) = 0$, all $t \neq s$ and have a constant unconditional variance in time, $E(\varepsilon_{it}^2) = \sigma_\varepsilon^2$. Defining a composite error term as $v_{it} = \mu_i + \varepsilon_{it}$; the variance of ν_{it} is $E(v_{it}^2) = \sigma_\mu^2 + \sigma_\varepsilon^2$ and $E(\nu_{it}\nu_{is}) = \sigma_\mu^2$, for all $t \neq s$.

The basic specification given in column 1 has been found after testing for several interaction effects and alternative specifications that will be pointed out as we discuss the main findings below. Among the alternatives, we have separate estimates for men and women and separate estimates by year of arrival. The latter distinction has been made, as before, to check if certain effects become more pronounced as immigrants have been in The Netherlands for a longer period of time. We have also estimated interaction effects with YSM; these were generally insignificant.

The gradient for age at arrival is fairly steep, with an annual growth rate of some 4%. The result is quite robust across specifications, but it drops if we estimate separately for later arrival cohort, suggesting that the disentanglement of age and YSM is less than perfect. There is a strong and very substantial positive direct effect of YSM. Quadratic effects of age and YSM have also been tested, but they were not significant. The effect of arrival year is fairly uniform for the first 3 years. The strong positive effect for the latest cohort may be a selectivity effect: These are refugees who can work right in their first year of arrival, which is quite unusual. Eliminating the dummies for arrival years has no effect on the estimates for age or YSM.

As the overall regression indicates, women earn about half of what comparable men earn, which is a striking difference. Several effects are essentially the same for men and women: age, YSM and marital status. The rankings by country are very similar, suggesting that country effects relate to real underlying differences in human capital that immigrants bring or the labour market views immigrant groups similarly. Just as for men, the coefficients on YSM do not differ significantly between countries. In fact, significance levels are even lower, and we can only conclude that in those early years after arrival, the speed of assimilation for women is identical across source countries. The only exception is Chinese women, with a strong positive effect.

The differences in status are not significant, except for AMAs when we split the sample by arrival time: In the youngest cohort, they are far behind, but in the oldest cohort, they have a premium of more than 30%. This is a fantastic race through the earnings distribution. The effects of time elapsed before obtaining status are quite interesting. Years spent as an undocumented worker add experience and increase pay (they may also signal that the individual in fact is not a refugee but an economic migrant, as a convinced refugee would start the application procedure right away; in that case, the additional pay may make up for initial low pay as an undocumented worker; see Hartog and Zorlu 1999). Conversely, years spent waiting for a status

reduce earnings, at about the same rate. Note that Statuswait covers time *before* YSM, whereas Undocyears covers years *parallel* with YSM. YSM starts at GBA registration, Statuswait is time spent in The Netherlands before GBA registration, and Undocyears is time spent since GBA registration. These are substantial rates: A year of undocumented work adds 13% to earnings on top of the benefits from YSM, and another year of waiting reduces earnings by 14%. The effects are primarily for men, as they are not significant for women. We have also tested for selection effects, by adding a dummy for immigrants who had returned by 2001. The coefficient is not significant, supporting our claim that in this sample, selective return migration is not an issue.

Married immigrants earn more than singles, and remarkably, on average, earnings are highest in The Hague, the seat of government. However, if we split between men and women, we see that men still earn most in The Hague, although women earn most in Amsterdam.

Education has an unexpected parabolic effect. Most coefficients are statistically significant. The returns peak for extended primary education. One might think that this reflects selectivity, as those with higher educations might be engaged in further education in the Dutch school system. However, the results of employment and benefit status in Section 4 (Table 5) do not lend much support to that interpretation. Interaction of education with YSM is insignificant for all levels of education. One might have thought that those with the highest education have the steepest time profiles because of complementarity between homeland education and the intensity and returns of investment in specific Dutch human capital (Duleep and Regets 1999). However, we did not find any significant interaction between schooling and YSM. We will return to such issues in the next section.

In columns 2 and 3, we present results separately for early and late arrivals (the earliest and the latest that we can meaningfully define; arrival is measured by year of IND registration; individuals may have been in the country before that, so we still have variation on YSM). The parabolic pattern of returns by education level is basically visible for the oldest and the youngest cohort, but precision is quite weak for the youngest. The oldest cohort have higher benefits from education than the youngest. It is quite remarkable that even for the oldest cohort, earnings drop for education levels beyond extended primary. There is no need to worry about effects of small sample size, as some 9% of the sample has higher education (for some countries, the percentage is well above 10, see Table 2). Also remarkable is the high pay for the least educated, some years primary, after 5 years in The Netherlands. Thus, benefits from education clearly increase with time spent in The Netherlands, but the pattern by level of education is surprising.

The parabolic effect of education that we found in the joint estimation is also visible in the results for men and women separately but with some differences. For men, returns to education behave like a step function: zero if basic education has not been completed, some 35% for primary and extended primary and some 20% for the higher levels. For women, a single peak stands out, a significant 41% at extended primary education.

The core result on education is a non-monotonic effect on earnings. Highest earnings are consistently found for immigrants with educations in the middle of the distribution. Most remarkable is the consistent drop in earnings for immigrants with education beyond secondary. How robust is this result?

In column 5 of Table 8, we have reported estimation results for the case where we drop all observations where information on education is missing. This has no effect: Whether we know education or not is immaterial for the estimation of the coefficients on the other variables. Covariances between education and other variables are not responsible for the result.

We have also made estimates with a selection on observations for reliability of the education variable, just as we did in the previous section for employment and benefit recipient status (Table 9). For ease of comparison, we copied the basic specification from Table 8.

The effect of selective observation by the Employment Service is remarkably small. The estimated coefficients differ somewhat between the full sample and the restricted sample used for model I, but in a qualitative sense, the conclusions are not affected. The coefficients on education are very similar, except for secondary vocational education. Immigrants with that education who visit the Employment Service are much more successful than an average immigrant with that education. Of course, we cannot say whether this is due to the positive influence of the Employment Service or to higher unobserved quality of those who visit. From inspecting results for models II and III, we can clearly conclude that our key conclusion on education survives. Immigrants with higher education do not earn

Table 9 Selecting on reliable measurement of education: selected coefficients of panel GLS random effect estimations of log weekly earnings, 1995–2000

	C1 from T8	Model I	Model II	Model III	Model IV	Model V
edu1_3 y	0.00	0.07	0.03	0.16		
edu4_5 y	0.21*	0.23*	0.17	0.10		
eduPrim	0.28***	0.27**	0.29**	0.40**		
eduPrim_ext	0.36***	0.32***	0.31**	0.25*		
eduSec_gen	0.31***	0.27***	0.27**	0.20		
eduSec_voc	0.21*	0.39***	0.41***	0.35**		
eduHigh_some	0.18	0.16	0.26	0.38**		
eduHigh	0.20**	0.27**	0.28**	0.26*		
edumissing	0.13*					
eduINDsec					0.06	
eduINDter					0.06	
eduSec_gen						2.91
eduSec_voc						−3.01
eduHigh_some						0.12
eduHigh						0.65
eduMissing						1.07
N-person	3,703	1,225	940	718	718	3,703
N-person-year	5,933	2,091	1,603	1,236	1,236	5,933

The models are estimated with the same variables as in Table 8 in addition to the education variables presented in Table 9 but their coefficients are not presented here.

Definitions of models: *C1 from T8*, basic specification (all wage earners): copied from first column of Table 8; *Model I*, sub-sample of individuals whose education is known; *Model II*, sub-sample of individuals whose education is known both in IND and CWI registers; *Model III*, sub-sample of individuals whose IND and CWI measures of education match perfectly; *Model IV*, sub-sample of individuals whose IND and CWI education match perfectly, three levels of education; *Model V*, IV estimates using all wage earners: CWI measure of education is used as an instrument for the IND measure of education

*$p<0.1$; **$p<0.05$; ***$p<0.01$

more than immigrants with lower education; education acquired at home does not pay off in the Dutch labour market. Under reliability restriction II, which restricts the sample to individuals whose education is known in both IND and CWI registers, the earnings levels for immigrants are identical for all education levels beyond some primary, with the exception of secondary vocational. Under reliability restriction III, which captures individuals whose IND and CWI measures match perfectly, there is equal pay for primary education, secondary vocational and some higher level education, with all other levels earning less. Model IV is even more outspoken: There is no earnings difference between immigrants with primary, secondary or tertiary education!

Results from instrumental variable regression are reported as model V. We used the CWI measure of education as an instrument for the IND measure of education. As we know that CWI education correlates with IND education and as we may assume that CWI education does not affect the disturbances in an earnings function that would include true education, CWI education is a good instrument (Wooldridge 2002, p. 83). We cannot maintain the same classification of education, as the number of instruments cannot be smaller than the number of instrumented variables. The instrumental variable estimation provides less precise coefficient estimates but confirms the earlier results that there is no monotonic increase in earnings with education level; higher educations do not lead to higher pay. These results are robust, no matter whether we use pooled OLS, random effect GLS or restriction to observations in 2000 only (the table only shows random effect GLS results). We should note, however, that instrumental variable estimation does not guarantee consistent estimates if the measurement error is not classical and that the bias cannot be signed (Kane et al. 1999).

We have considered estimation of earnings functions corrected for participation using Heckman's two-step procedure. A priori, we had reservations because not many variables are available and credible exclusion restrictions are hard to determine. We estimated a wage equation for the pooled sample with correction of standard errors for repeated observations and separate for 2000 only. If the wage equation includes education and country of origin, we get unconvincing results no matter how we specify the participation equation. In particular, the effect of YSM is negative, and the dummy for women gets a positive coefficient. We decided not to pursue this approach.

6 Dip and catch-up: testing Duleep and Regets

Duleep and Regets (1999) have used the human capital model to derive testable predictions for the "dip and catch-up" model. In this model, immigrants start out at an economic disadvantage relative to natives but with increasing duration in the destination country they may catch up, with faster growth rates in wages and employment probability. Newly arrived immigrants have lower opportunity cost of human capital investment than natives because of their wage dip upon arrival. They will also have higher returns to the extent that investments in the destination country human capital increase the value of their home country human capital. Thus, they will invest more and have faster earnings growth (and thus will catch-up). Skill transferability between the home country and the destination country is an important variable, as this will affect the magnitude of the initial dip. The differences in wages and employment between source countries may be related to differences in skill

transferability, but we do not have any additional data (e.g. on schooling systems), and we cannot test this theory. We may note, however, that we have not found significant interaction effects between education and country of origin, between YSM and education or between YSM and source country in wage regressions. These issues need further research.
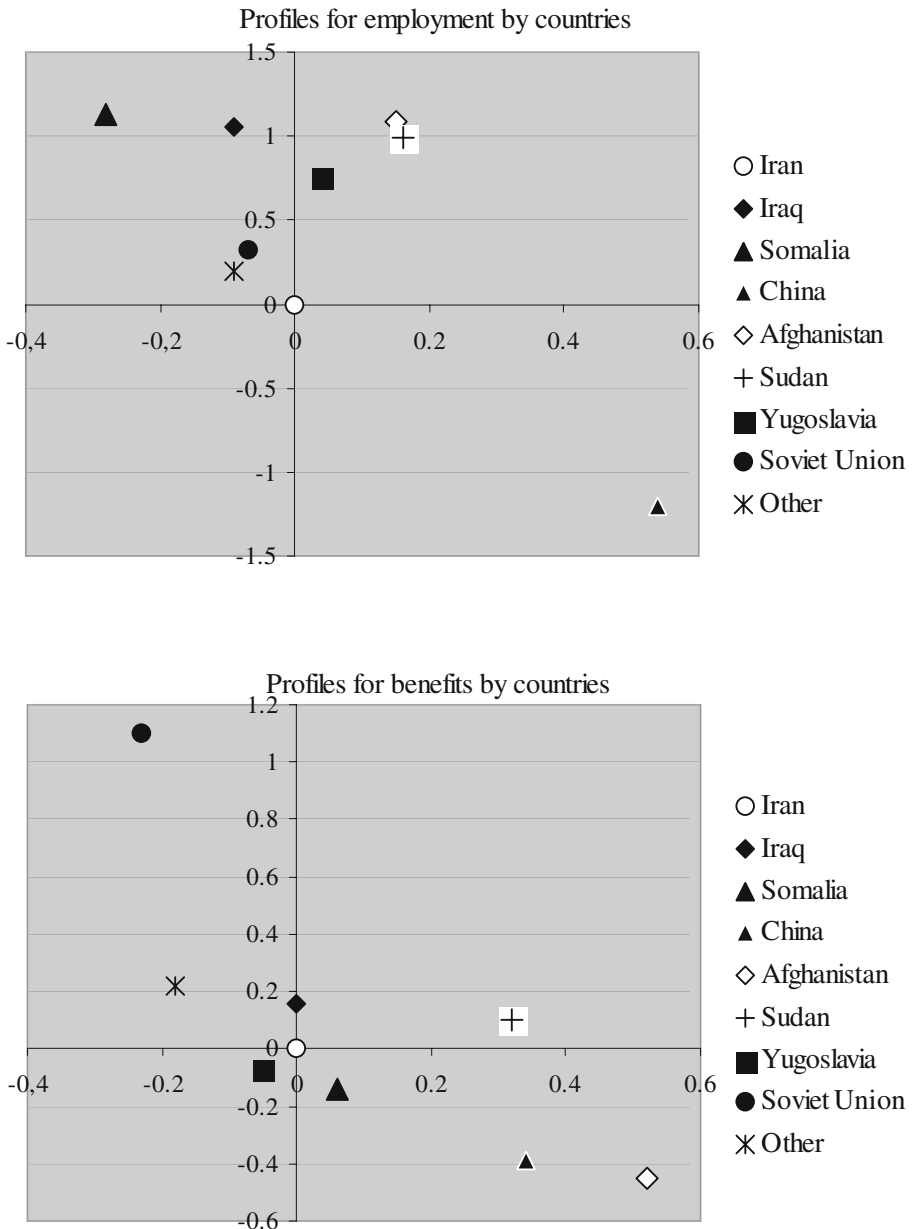

Fig. 3 Country intercepts and country slopes on YSM

The core prediction of a negative relation between initial dip and the slope on YSM is generally supported. In Fig. 3, we plot the dummy coefficient for a source country against the interaction coefficient for that country with YSM, from three regressions: the multinomial logits for employment and benefit status (Table 5) and the basic reference regression for wages in Table 8 (model I). This means that we test the prediction by comparing immigrants from different countries rather than comparing immigrants with natives. For employment probability, the negative relationship between intercept and slope is strong dependent on the observations for Somalia and China. Both for benefits and wages, the marginal effect of length of stay in The Netherlands is generally larger for immigrants from countries with a smaller country intercept. Note that this is not a necessary mechanical relationship. Although a larger gap with natives indicates greater potential for growth, there is no need for this potential to be realised.

The prediction appears to hold also for the probability to receive social benefits. This suggests that the hypothesis on human capital might also apply to the investment in the social capital: getting to know your way around the institutions.

7 Conclusion and possible explanations

Our key finding is that for refugees, higher educations acquired at home generally do not pay off during the first 5 years in the Dutch labour market. Although remarkable, the outcome matches observations of persons active in refugee circles. We discussed our results with the immigration department of the Ministry of Justice and with Vluchtelingenwerk Nederland, a foundation that supports newly arrived refugees. They were not surprised. The result may be explained in several ways. One intervening variable may be language skills. It may very well be that for many of the occupations associated with higher educations, understanding the Dutch language is vital, much more so than for lower levels of education. One can do cleaning work, construction work and much manufacturing work without good fluency in Dutch, as the results for Turkish and Moroccan immigrants in The Netherlands testify. One cannot be a physician without a high level of competence in Dutch. As we have no information on language proficiency, we cannot test this, but such a hypothesis is clearly supported by Berman et al. (2003) for Israel. It would be quite informative to observe jobs that immigrants hold before and after immigration, but such information is not available.

A related explanation may be certification. Several occupations that require high levels of education also require certification in the destination country. Even if one were fluent in Dutch, a qualified physician would not be allowed to take up his profession without obtaining new professional qualification in The Netherlands. Certification may have elements of discrimination and job protection but may also have a basis in country-specific required skills.[5] Of course, even without certification, there may be plain discrimination. Without further data, however, we cannot assess the empirical importance of these explanations.

---

[5] At the Mannheim presentation of this paper, someone remarked that his taxi driver to Stockholm airport had been a former Iraqi army general. Perhaps his skill might have been transferable, but his occupation certainly was not.

Another explanation may be possible differences in health condition and true immigration motive. Refugees usually experience violence in their home country and may carry health-affecting consequences of repression. Moreover, there is some doubt that every asylum seeker is a (political) refugee. Because legal immigration from developing countries is highly restrictive, some (economic) immigrants try to enter The Netherlands via the asylum procedure. If political engagement is correlated with higher education levels, the population of economic immigrants who applied and secured a refugee position is mostly from the lower end of the educational distribution. Therefore, lower skilled refugees, if mainly economic immigrants, may be successful in the labour market, whereas the value added of higher skills of real refugees might be offset by health problems and traumatic experiences that hamper their integration in a new society.

Appendix A

Sample characteristics of the immigrant panel

Table 10 Descriptive statistics, 2000

	Mean	Std. Dev.
Age	28.47	11.31
Woman	0.37	0.48
Married	0.40	0.49
YSM	1.97	1.48
Undocyears	0.92	0.83
Statuswait	0.07	0.50
A status	0.20	0.40
AMA	0.06	0.24
Naturalised	0.16	0.37
Returned	0.02	0.14
Amsterdam	0.06	0.23
Rotterdam	0.05	0.22
The Hague	0.05	0.22
Utrecht	0.01	0.12
Employed	0.32	0.47
Benefits	0.35	0.48
Weekly wages	183.60	174.88
N	13,436	

Table 11 Years of education by source country (mean, standard deviation)

	IND			CWI		
	Mean	Std. Dev.	Frequency	Mean	Std. Dev.	Frequency
Iran	10.89	3.76	387	10.27	3.02	296
Iraq	10.33	4.79	1546	10.19	3.67	1,292
Somalia	5.72	4.75	602	8.38	2.64	354
China	6.14	3.01	316	8.04	2.22	159
Afghanistan	9.95	5.13	1,186	9.46	3.52	1,780
Sudan	10.77	4.93	364	10.18	3.36	246
Yugoslavia	9.06	3.92	780	9.71	2.88	516
Soviet Union	10.24	4.29	506	9.79	3.63	328
Other	7.77	4.50	1,757	9.16	2.95	1,039
Total	9.02	4.82	7,444	9.57	3.35	6,016

Appendix B

Variable definitions

Observations restricted to individuals aged 15–59

Arrival.year (instroom): year of registration IND

Settlement.year (vestiging): year of registration GBA

Age: age at arrival in The Netherlands, either at IND or GBA registration

YSM: years since migration; years elapsed since registration GBA

Statuswait: settlement.year minus arrival.year, if positive, zero otherwise (year of GBA registration minus year of IND registration), hence time spent in refugee homes waiting for a decision on the application

Undocyears: arrival.year minus settlement.year, if positive, zero otherwise (year of IND registration minus year of GBA registration); this applies when immigrants settled in the Netherlands without residential permission and without applying, undocumented immigrants could register at GBA without any sanction

Table 12 Years of education as registered by IND and coded by ITS and as registered by CWI (Employment Office): average years added by us

IND		CWI	
	Years of education		Years of education
No education	0	Basic	6
1–3 year Primary	3	Lower	10
4–5 year Primary	5	Intermediary	12
Primary	8	Higher vocational	15
Extended Primary	9	University	16
Secondary, general	11		
Secondary, vocational	12		
Some Tertiary	14		
Tertiary	16		

A status: permanent residential permission; date of granting status unknown
AMA: independent refugee not older than 18 at arrival
Preliminary status: not officially recognised as refugee, but admitted until home country is declared safe
Naturalised: obtained Dutch citizenship in 2001
Returned: emigrated or administratively removed in 2001
Married: individual had marital status when arriving.

References

Altonji JG, Blank RM (1999) Race and gender in the labor market. In: Ashenfelter O, Card D (eds) Handbook of labor economics, vol. 3, part C. Elsevier, Amsterdam, pp 3143–3259

Battistin E, Sianesi B (2006) Misreported schooling and returns to education: evidence from the UK Institute for Fiscal Studies, London. Working Papers, CWP07/06

Berman E, Lang K, Siniver E (2003) Language-skill complementarity: returns to immigrant language acquisition. Labour Econ 10(3):265–290

Bevelander P, Nielsen H (2001) Declining employment success of immigrant males in Sweden: Observed or unobserved characteristics? J Popul Econ 14(3):455–471

Bratsberg B, Ragan JF Jr (2002) The impact of host-country schooling on earnings: a study of male immigrants in the United States. J Hum Resour 37(1):63–105

Card D (1999) The causal effect of education on earnings. In: Ashenfelter O, Card D (eds) Handbook of Labor Economics, vol. 3, part A. Elsevier, Amsterdam, pp 1801–1863

Chiswick BR, Miller PW (1994) The determinants of post-immigration investments in education. Econ Educ Rev 13(2):163–177

Chiswick BR, Miller PW (2003) The complementarity of language and other human capital: Immigrant earnings in Canada. Econ Educ Rev 22(5):469–480

Chiswick BR, Lee YL, Miller PW (2005) Immigrant earnings: a longitudinal analysis. Rev Income Wealth 51(4):485–503

Cortes K (2004) Are refugees different from economic immigrants? Some empirical evidence on the heterogeneity of immigrant groups in the United States. Rev Econ Stat 86(2):465–480

Cramer JS (2003) Logit models from economics and other fields. Cambridge Univ. Press, Cambridge, UK

Duleep H, Regets M (1999) Immigrants and human capital investment. Am Econ Rev 89(2):186–191 (Papers and Proceedings)

Friedberg R (2000) You can't take it with you? Immigrant assimilation and the portability of human capital. J Labor Econ 18(2):221–251

Hartog J, Zorlu A (1999) Turkish clothing in Amsterdam: the rise and fall of a perfectly competitive labour market. De Economist 147(2):151–181

Husted L, Nielsen HS, Rosholm M, Smith N (2001) Employment and wage assimilation of male first-generation immigrants in Denmark. Int J Manpow 22(1/2):39–68

Kane TJ, Rouse CE, Staiger D (1999) Estimating returns to schooling when schooling is misreported. NBER Working Paper 7235

Kee P (1993) The economic status of male immigrants in the Netherlands. Ph.D. thesis, University of Amsterdam, Amsterdam

Nekby L (2002) How long does it take to integrate? Employment convergence of immigrants and natives in Sweden. FIEF Working Paper Series no. 185, Stockholm

Wooldridge J (2002) Econometric analysis of cross section and panel data. MIT, Cambridge