DISCUSSION PAPER SERIES

IZA DP No. 8627

Immigrants' Wage Growth and Selective Out-Migration

Govert E. Bijwaard
Jackline Wahba

November 2014

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Immigrants' Wage Growth and Selective Out-Migration

Govert E. Bijwaard
Netherlands Interdisciplinary Demographic Institute (NIDI)
and IZA

Jackline Wahba
University of Southampton
and IZA

Discussion Paper No. 8627
November 2014

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 8627
November 2014

ABSTRACT

Immigrants' Wage Growth and Selective Out-Migration*

This paper examines immigrant wage growth taking into account selective out-migration using administrative data from the Netherlands. We also take into account the potential endogeneity of the immigrants' labor supply and their out-migration decisions on their earning profiles using a correlated competing risk model, but we also use standard estimations as done in previous literature. We distinguish between two types of migrants: labor and family migrants given their different labor market and out-migration behavior. We find that simple models lead to biased estimates of the wage growth of immigrants. Controlling for the selective out-migration and endogeneity of labor supply, we find that labor out-migrants are positively selected but family out-migrants are negatively selected. Furthermore, the findings underscore the importance of taking into account the endogeneity of labor supply and out-migration when estimating immigrants' wage growth.

JEL Classification: F22, J61, C41

Keywords: migration dynamics, labor market transitions, competing risks, immigrant assimilation, income growth

Corresponding author:

Govert E. Bijwaard
Netherlands Interdisciplinary Demographic Institute
(NIDI-KNAW/University of Groningen)
PO Box 11650
2502 AR The Hague
The Netherlands
E-mail: bijwaard@nidi.nl

* Financial support from the NORFACE research programme on Migration in Europe – Social, Economic, Cultural and Policy Dynamics and the Economic and Social Research Council [RES-167-25-0678] is gratefully acknowledged. We thank Statistics Netherlands, Han Nicolaas in particular, for data support.

1 Introduction

How do immigrants fare in the host country? Many studies are concerned with whether immigrants assimilate in the labor market over time. Not surprisingly, this is driven by the public interest in whether immigrants are a burden on the fiscal system or if they are contributors to the welfare state. In particular immigrants' earning profiles are commonly used as a measure of their success and contribution to the economy. A sizable literature has examined immigrant performance in the host country focusing on immigrant earnings growth. This literature has evolved enough to deal with several empirical challenges namely, disentangling the effect of migrant age at arrival, cohort effects and time effects. Hence, the earlier studies based on cross-sectional data e.g. Chiswick (1978) and Borjas (1985), find that immigrants experience rapid wage growth over time. In contrast, those using longitudinal or panel data typically find smaller immigrant' wage growth (Hu, 2000; Lubotsky, 2007). However, a well-known and long-recognized problem within this literature is the nonrandom out-migration bias. The major challenge faced by these studies in tackling out-migration selectivity is the absence of administrative data collected by immigration authorities, where information on out-migration is observed rather than inferred from attrition from survey samples (e.g. Pischke 1992). To some extent, recent papers by Abramitzky et al (2012) and Biavaschi (2012), try to circumvent this challenge by using repeated census data of both the sending and the host country. Yet, again this is problematic if not all migrants (returnees) are identified in the host (sending) country census and different definitions of education are used in the sending and the host country.

This paper aims to address this gap in the literature by examining immigrant wage growth taking into account the out-migration selection. We use administrative data from the Netherlands, where we observe all immigrants who have entered the country between 1999 and 2007, and their motive for migration: whether for labor migration or otherwise, the timing of return in case of out-migration, and the exact detailed (monthly) information on their labor market status and income since their arrival. This enables us to better understand how selective out-migration influences immigrant wage growth.

We also address an additional challenge that has been overlooked in this literature; namely, that out-migration might be related to the level of immigrant earnings: migrants leave because of their wage levels or wage growth. On one hand it could be the unsuccessful ones who leave because of a lack of employment and negative wage growth, or it could be the successful migrants who, having experienced positive wage growth and having reached their target savings, decide to leave. Hence, our main contribution is to address this potential endogeneity in the out-migration selection. We take

into account the correlation between three important factors which has not been considered before: the labor market status of the migrant since arrival, her wages in the country of immigration and the out-migration decision as they can all affect the estimated wage growth of current immigrants. We do not assume as previously done in the literature, and noted by Dustmann and Görlach (2014), that immigrants' decisions on labor supply are *not* determined by their out-migration plans.

Furthermore, different types of migrants, for example, labor migrants and family migrants, behave differently in the labor market and when it comes to out-migration. Census data and standard panel surveys do not distinguish between migrant types. Hence, previous studies have not been able to distinguish between the behavior of different types of migrants, in particular labor versus family migrants. In the Netherlands, as in other European countries, family migrants represent a higher proportion of total migrants. For example, about 23% of all non-Dutch immigrants, in the age group 18 - 64, are labor migrants and 25% are family migrants. Our data allow us to distinguish between those two potentially different types of migrants who might experience different wage trajectories and different patterns of out-migration selection and if aggregated as one group might potentially create biased results.

There are several reasons for the importance of accurately measuring immigrant earnings growth, which is usually seen as a measure of their success and assimilation. Many fear that immigrants will be a drain on the welfare system if they are not successful in the host country. Secondly, the economic status of immigrants is an important source of income inequality and poverty persistence in host countries.

Our empirical strategy is to adopt a correlated multistate model (CCRM) to estimate immigrant earnings growth and address the endogeneity of wages, labor market status and out-migration selection taking into account the unobserved heterogeneity among immigrants. We first follow the previous literature and adopt a similar approach by estimating simple Mincerian wage equations (and Heckman corrected wage equations) using standard types of data to compare those estimates to those obtained from our preferred model (CCRM). We use our data to construct the three main types of data used by previous studies namely (i) survey cross sectional data, (ii) repeated cross section data and (iii) panel data. We limit our analysis to the first migration spell; i.e. we do not include repeat migrants who left the Netherlands and then came back. Our findings suggest that ignoring out-migration selectivity biases the estimates of immigrants' wage growth but in a non-monotonic way: it overestimates wage growth in the early years of migration and underestimates it in later years. Furthermore, our results highlight the importance of taking into account the endogeneity of the labor supply and out-migration

when estimating immigrant earnings profiles.

The outline of the paper is as follows. In Section 2, we present the data and discuss the main characteristics of recent migration to the Netherlands, before summarizing the main theoretical and empirical literature on immigrant earnings growth in Section 3. Section 4 provides the results of estimating immigrant earnings as previously done in the literature by using similar (constructed) data. Section 4.5 introduces our correlated multistate model (CCRM) which takes the endogeneity of wages, labor market status and out-migration selection into account. Section 5 distinguishes between various migrant groups: it presents the estimates for labor vs. family migrants, compares the findings by main region and country of origin and finally provides the estimates of immigrant earnings profiles by initial income. The last section concludes.

2 Administrative panel data on the population of immigrants to the Netherlands

All legal immigration by non-Dutch citizens to the Netherlands is registered in the Central Register of Foreigners (Centraal Register Vreemdelingen, CRV), using information from the Immigration Police (Vreemdelingen Politie) and the Immigration and Naturalization Service (Immigratie en Naturalisatie Dienst, IND). It is mandatory for every immigrant to notify the local population registrar immediately on arrival in the Netherlands if he or she intends to stay for at least two-thirds of the forthcoming six months. The data comprise the entire *population* of immigrants who entered during our observation window of 1999 - 2007, and after merging it with other administrative registers we obtain a panel.

In addition to the date of entry and exit, the CRV also records the migration motive of the individual. Either the motive is coded according to the visa status of the immigrant, or the immigrant reports the motive on registration in the population register. Statistics Netherlands distinguishes between the following motives: labor-migrants, family migrants, student immigrants, asylum seekers (and refugees), and immigrants with other motives. Note that EU-citizens are required to register in The Netherlands, just as natives are. See Bijwaard (2010) for an extensive descriptive analysis of the various migration motives. We focus on the two main migration motives: labor and family migrants. In particular, about 23% of all non-Dutch immigrants in the age group 18 - 64 are labor migrants and 25% are family migrants. As it is possible that the official migration motive does not always match with the true intention of the migrants, we further require that labor migrants are employed in the Netherlands within three months of their entry.

The immigration register is linked by Statistics Netherlands to the Municipal Population Registry

(Gemeentelijke Basisadministratie, GBA) and to the Dutch Social Statistical Database (SSD). The GBA contains basic demographic characteristics of the migrants, such as age, gender, marital status and country of origin. From the SSD we have monthly information on the labor market status, income, industry sector, housing and household situation. To capture the economic conditions of the country of origin, we use annual GDP per capita and GDP growth rate by country of origin from the World Bank, World Development Indicators.

2.1 Descriptive statistics

First, we provide an overview of our data which are comprised of 194,775 immigrants: 92,893 labor immigrant and 101,882 family immigrant. Almost 53% are women, 64% are married and 70% are younger than 35 years of age. Table 1 shows immigrant income growth over time spent in the host country. Column 1 shows the average income for all the immigrants, whilst Column 2 shows the average income for migrants with positive income only; i.e. exclude migrants with zero income. Column 3 displays our group of interest, namely employed migrants (with positive income) only. The first row shows average income at time of arrival in the Netherlands. The second row presents the last observed income which includes the last income of those who left. Row three displays the income at the end of our observation window 1/1/2008, and thus captures stayers only, followed by the average income for years since migration (YSM). There are a number of interesting observations to note. First, the average income for all migrants is lower at the time of entry than at the time of last observed or at the end of our observation window. However, for those migrants employed with positive earnings, their average income before returning (last observation) tends to be higher than at the time of entry or at the end of the observation window suggesting that high earners are likely to leave hence also the average income of remaining migrants seems to decline over time (with YSM). It is important to note that those descriptive statistics confound both the cohort and time effects.

3 Theoretical and empirical background

The success of immigrants in the labor market is commonly measured by the growth in their earnings. A key issue in the debate on migration is whether immigrants contribute to or take from the welfare system. Typically immigrants with low wages are net receivers and those with high wages are net contributors. In this respect the evolution of the income of the immigrants over their migration cycle is the most relevant indicator of the economic performance.

Immigrants, when they first arrive in the new host country, often lack country-specific human

Table 1: Descriptive statistics income

	Average	Average (>0)[a]	Average (>0, Emp)[b]
entry	€1375	€3198	€3238
last observation	€1630	€3122	€3451
observation end	€1988	€2985	€3311
YSM			
0–1 year	€2571	€3741	€3791
1–2 year	€2395	€3674	€3351
2–3 year	€1949	€3279	€3594
3–4 year	€1989	€3133	€3536
4–5 year	€1802	€2813	€3220
5–6 year	€1626	€2456	€2798
6–7 year	€1817	€2631	€3006
7–8 year	€1791	€2543	€2849
8–9 year	€1796	€2525	€2929

[a] Only for migrants with positive income.
[b] Only for employed migrants with positive income.

capital—such as knowledge about the labor market, and language skills. More time spent in the host country is assumed to lead to more human capital investment that is likely to increase the productivity of immigrants and enable them to progress up the earnings ladder. The effect of years since migration is expected to have a positive effect on immigrant earnings, given such investments in host country specific human capital. This investment process explains the relatively rapid rates of growth in immigrant earnings observed in cross section studies: The seminal work, Chiswick (1978), provided the first insights about that relationship using cross-section estimates, assuming though that the unobserved characteristics of the immigrants are stable over time. However, the positive impact of the years-since-migration variable in cross-sectional earnings equations captures both the quality of earlier immigrant cohorts as well as the increase in the host specific human capital. Subsequent work by Borjas (1985) and LaLonde and Topel (1992) using repeated cross sections are able to identify the cohort effect but assume that the time effects are the same for natives and immigrants. Finally, repeated census data are still problematic as the estimate of the cohort fixed effects is a mixture of earnings of those who stay and those who leave.

The use of longitudinal data on earnings histories has helped to overcome some of the limitations of the previous literature on immigrant earnings assimilation (e.g. Hu (2000) and Lubotsky (2007)). Nonetheless, in assessing the economic performance of immigrants it is often unaccounted for that many migrants leave the host country, or if acknowledged little is done about it given the lack of adequate data. Dustmann and Görlach (2014) provide an excellent review of the potential biases and identifying assumptions associated with the various types of data adopted in the literature and their

impacts on the estimates of immigrant earnings profiles. The level of out-migration has been high for both the US and many European countries. Jasso and Rosenzweig (1982) report that between twenty and fifty percent of legal immigrants re-emigrated from US in the 1970s. Bijwaard (2010) reports return probabilities between 25% and 50% within five years since immigration, depending on the migration motive, for migrants entering the Netherlands in 1995-2003. Similar numbers have been found for other Western European countries, see e.g. Edin et al. (2000) for Sweden, Jensen and Pedersen (2007) for Denmark and Dustmann (1995) for Germany. Despite the knowledge that many migrations are temporary the majority of the literature on migration still (implicitly) assumes migrations are permanent (see e.g. Chiswick 1978, Massey et al. 1993 and Borjas 1999). When many migration decisions are temporary and the migrants that re-migrate are more (less) successful than the migrants that remain in the country, ignoring such selective out-migration will under(over) state the economic performance of the original immigrants who have stayed.

One important aspect affecting immigrant behavior is that economic decisions related to the labor market are usually made in conjunction with return migration decisions. Thus, the out-migration might be endogenously determined with the labor market status and income growth. This is another empirical challenge that has not been addressed in the wage growth literature. To our knowledge the only paper that studies the influence of individual labor market changes on the return decision is Bijwaard and Wahba (2014) which has a different focus, namely whether the high or low income immigrants leave faster. A related paper by Bijwaard et al. (2014) uses a timing of events model to control for the endogeneity of labor market changes to examine the impact of unemployment on return migration. They find that unemployment induces labor migrants to return and re-employment makes the migrants more prone to stay. However, neither of these papers study wage growth. Indeed, our current paper's contribution is in adopting a correlated multistate model (CCRM) to estimate immigrant earnings growth and addressing the endogeneity of wages, labor market status and out-migration selection taking into account the unobserved heterogeneity among immigrants. In order to show the significance of the type of data used on the estimates of immigrant earnings profile, we follow the standard literature and treat our data as single cross section, repeated cross section and a panel before using the administrative longitudinal data of the full immigrant population at our disposable.

4 Estimating wage growth of immigrants

In Section 4.1 we follow the earlier literature and construct three types of data all of which ignore out-migration. First, we assume that we observe a 'survey' conducted at 1/1/2008 containing all the recent

migrants who entered between 1/1/1999 and 31/12/2007 and who are in the country on 1/1/2008. Assuming a survey implies that we only observe their current (monthly) income, employment status and all other relevant individual characteristics. We know the entry date and we derive from that the years-since-migration variable. Based on these data we obtain OLS estimates and estimates from a Heckman selection model that corrects for self-selection into employment. In Section 4.2 we assume we observe only those immigrants who are in the country on January 1$^{\text{st}}$ of each year, 1999-2008. These are repeated cross section data. In this setting we know when the migrant arrived, but ignore that we have repeated observations of the same individual. Based on these data we estimate an OLS, and an OLS Heckman selection model. In Section 4.3 we assume that we observe for all the stayers, those immigrants still in the country at the end of the observation window, the monthly employment and income history (and the history of other relevant migrant characteristics). Based on these panel data we obtain the fixed effects (FE) estimates for migrants. In Section 4.4 we use the full history of the stayers and estimate a correlated multistate model with three labor market states. In Section 4.5 we include all the migrants and extend the correlated multistate model with an absorbing living-abroad state to capture selective out-migration.

4.1 Survey data

Our starting point is replicating the standard methodology used by previous studies e.g. by Chiswick (1978) and use a (constructed) cross section dataset. It is standard practice in the literature to estimate the wage growth of immigrants considering only those migrants who are still in the country by the end of the observation window which in our case is 1/1/2008. These migrants have stayed for maximum 9 years in the country. The standard wage equation is:

$$\ln W = \theta_0 + \sum_{h=1}^{8} \theta_h YSM_h + \theta X + u_w, \tag{1}$$

where YSM_h is years-since-migration dummies and X is other control variables that influence the wage, like gender, sector, and age. To control for the cohort effect we use the unemployment rate in the Netherlands at time of entry. We assume that the errors are normally distributed, $u_w \sim \mathcal{N}(0, \sigma^2)$.

Of course, we only observe wages for those migrants who are employed at the end of the observation window. A standard procedure to correct for the (possible) selection into employment of the migrant is to use a Heckman selection model in which we add the selection equation

$$E^* = \gamma_x X + \gamma_z Z + u_e. \tag{2}$$

A migrant is employed, $E = 1$, when $E^* > 0$ and $u_e \sim \mathcal{N}(0, 1)$ and $\text{corr}(u_w, u_e) = \rho$. In the selection equation we add the instrumental variables married, divorced and number of children.

First, we assume that we only have a 'survey' conducted at the end of our observation window. This survey contains information on all migrants who are still in the country at that moment. We ignore the longitudinal nature of our data and pretend we only observe the information at 1/1/2008. However, for every immigrant we do observe the day of entry, which allows us to calculate the years-since-migration. Table 2 shows the OLS and Heckman selection estimates based on the assumption that we only have such survey data available. In both cases, years-since-migration (YSM) has a significant positive impact and, on average, the coefficients are smaller once we control for the participation/employment decision. Overall, those results suggest that immigrant wages have grown over time, but not monotonically.

Table 2: Impact of years-since migration (YSM) on log-wage of migrants, survey data

	OLS	Heckman
YSM2	0.074^{**}	0.074^{**}
	(0.013)	(0.013)
YSM3	0.152^{**}	0.152^{**}
	(0.020)	(0.020)
YSM4	0.152^{**}	0.145^{**}
	(0.022)	(0.022)
YSM5	0.125^{**}	0.118^{**}
	(0.017)	(0.018)
YSM6	0.078^{**}	0.076^{**}
	(0.012)	(0.012)
YSM7	0.076^{**}	0.076^{**}
	(0.011)	(0.011)
YSM8	0.100^{**}	0.096^{**}
	(0.011)	(0.011)
YSM9	0.131^{**}	0.125^{**}
	(0.013)	(0.013)

$^{+}p < 0.05$ and $^{**}p < 0.01$

Of course, using cross sectional survey data is problematic as we can not distinguish between the cohort effects and the individual fixed effects. Next we turn to data that have been collected annually over the whole period, ignoring repeated observations of the same individual.

4.2 Repeated cross section data

Data are often collected at fixed periods, e.g. census data, as used by Borjas (1985) and LaLonde and Topel (1992) for example. When no personal identification is available the data are in the form of a repeated cross section. We assume that the entry date is available. Thus, for a migrant who arrived mid- 1999 and stayed in the country till the end of the observation window we have nine independent observations on January 1st of each year, 2000-2008. Again we estimate an OLS model in (1), and a Heckman selection model to correct for selection into employment (i.e. positive income).

Table 3 shows these estimates based on our constructed repeated cross section data. In both models the repeated cross section estimates are above the survey data estimates. The OLS results seem to indicate that, after an initial wage increase, wages decline with the length of time in the country. This is probably caused by an increasing number of migrants who are not employed. Correction for the selection into employment (Heckman) decreases the estimated impact of years since migration. Note that the results of the OLS and the Heckman correction model are very close indicating only a minor effect of selection into employment on wage growth. However, the curves are still rather flat (see also Figure 1). Thus, these results suggest only a small initial wage increase in immigrant wages with the time spent in the country.

Table 3: Impact of years-since migration (YSM) on log-wage of migrants, repeated cross section

	OLS	Heckman
YSM2	0.140**	0.138**
	(0.006)	(0.006)
YSM3	0.187**	0.181**
	(0.007)	(0.007)
YSM4	0.193**	0.183**
	(0.007)	(0.007)
YSM5	0.200**	0.191**
	(0.008)	(0.008)
YSM6	0.183**	0.174**
	(0.009)	(0.009)
YSM7	0.180**	0.171**
	(0.010)	(0.010)
YSM8	0.153**	0.142**
	(0.012)	(0.012)
YSM9	0.134**	0.121**
	(0.016)	(0.016)

$^{+}p < 0.05$ and $^{**}p < 0.01$

4.3 Panel stock data

Next we assume that the data are from a retrospective panel, based on the stock of migrants still in the country at the end of the observation period. For example, Pischke (1992) and Borjas (1989) among others used such longitudinal data. For these migrants we observe their full monthly employment and income history (and the history of other relevant migrant characteristics). Such data allows us to estimate both random effects (RE) and fixed effects (FE) panel models that exploit the full information available for every migrant. Panel data have a distinct advantage, namely allowing for several potential biases due to correlated individual effects to be controlled for; such biases arise from the cohort effects, and the role of age at entry. Yet panel data cannot control for out-migration selection if attrition can not be disentangled from remigration. Panel attrition will bias the estimation results if the probability of leaving the sample is systematically linked to labor market status and wages. The estimates in Table 4 (and Figure 1) indicate that the use of panel data suggests a higher impact of YSM on wage growth than using (repeated) cross-section data.

Table 4: Impact of years-since migration (YSM) on log-wage of migrants, panel stock data

	Random effects	Fixed effects
YSM2	0.304**	0.165**
	(0.002)	(0.001)
YSM3	0.414**	0.267**
	(0.003)	(0.002)
YSM4	0.457**	0.332**
	(0.004)	(0.002)
YSM5	0.462**	0.378**
	(0.005)	(0.003)
YSM6	0.456**	0.415**
	(0.006)	(0.003)
YSM7	0.459**	0.442**
	(0.007)	(0.004)
YSM8	0.465**	0.470**
	(0.008)	(0.004)
YSM9	0.528**	0.528**
	(0.017)	(0.007)

$^{**}p < 0.01$

4.4 Stock data with full labor market history

The wage growth of an immigrant depends heavily on the labor market performance of the migrant in the host country. Migrants who are employed all the time have more likely to get higher wages than migrants who are unemployed or non-participating for a while. It is most likely that the transitions among these states and the wage growth depend on both observed and unobserved factors. So in this section, we diverge from the previous literature by showing the significance of the labor market history. Hence, we restrict the analysis to the immigrants who remain till the end of our observation window, but we observe their full labor market history, which we call stock data with full labor market history. An immigrant potentially faces several labor market transitions during his stay in the host country and multiple durations in the different states. Thus, we consider three labor market states:

1. Employed in the host country;
2. Unemployed and receiving benefits in the host country;
3. Out of the labor market (includes both unemployed but not receiving benefits and non labor marker participants) in the host country;

These states are mutually exclusive and exhaust all possible destinations. A migrant may leave a state $j = 1, \ldots, 3$ for any of the other destination states, i.e. for $j = 1$ the destination states are $k = 2, 3$, for $j = 2$ the destination states are $k = 1, 3$ etc. We view the migrant behavior as a semi-Markov process with individuals moving between the three states.

We use a competing risks hazard model for each origin-destination pair. We define the random variables T_{jk} that describe the time since entry in state j for a transition from j to k. We assume a mixed proportional hazard model for which the intensity for the transition from j to k is:

$$\lambda_{jk}(t|\overline{X}_{jk}(t), V_{jk}) = \lambda_{0jk}(t) \exp\big(\beta'_{jk} X_{jk}(t) + V_{jk}\big) \tag{3}$$

where $\overline{X}_{jk}(t) = \{X_{jk}(s)|0 \leq s \leq t\}$ is the sample path of the observed characteristics up to time t, which is, without loss of generality, assumed to be continuous. The unobserved heterogeneity V_{jk} also enters the intensity multiplicatively. We assume that the path of the observed characteristics is independent of the unobserved heterogeneity. The positive function $\lambda_{0jk}(t)$ is the baseline intensity and we assume that it is piecewise constant on H intervals[1], i.e. $\lambda_{0jk}(t) = \sum_{h=1}^{H} e^{\alpha_{jkh}} I_h(t)$ with

[1]It is not necessary that each baseline intensity has the same duration. Here H is the total number of intervals considered. If, for the transition from j to k, the baseline intensity remains the same in $I_h(t)$ and $I_{h+1}(t)$, we have $\alpha_{jkh} = \alpha_{jkh+1}$.

$I_h(t) = I(t_{h-1} \leq t < t_h)$ and $t_0 = 0$, $t_H = \infty$. Any duration dependence can be approximated arbitrarily closely by increasing the number of intervals. The integrated intensity for a transition from j to k at duration t is (conditional on V)

$$\Lambda_{jk}(t|\overline{X}_{jk}(t), V_{jk}) = \sum_{h=1}^{H} e^{\alpha_{jkh}+\beta_{jk}X_h+V_{jk}}\big(t_h - t_{h-1}\big)J_h(t) + \sum_{h=1}^{H} e^{\alpha_{jkh}+\beta_{jk}X_h+V_{jk}}\big(t - t_{h-1}\big)I_h(t) \quad (4)$$

with $J_h(t) = I(t > t_h)$ and we assume that any change in the time-varying components of X only occurs at discrete times and that the H intervals also capture these changes. Thus, x_h is the value of x in interval $[t_{h-1}, t_h)$. For identification we assume the baseline hazard is one in the first interval, i.e. $\alpha_{jk1} = 0$.

For each origin state, only the smallest of T_{jk} durations $\tilde{T}_j = \min_k T_{jk}$ and the corresponding actual transition destination is observed. The other durations are censored, in the sense that all is known that all we know is that their realizations exceed $\tilde{T}_j$. If for individual i we observe M_{ijk}, for j to k transition spells, at sojourn times $t_1, \ldots, t_M$, then the likelihood contribution of these M_{ijk} transitions is:

$$L_{jk}(V) = \prod_{m=1}^{M_{ijk}} \lambda_{jk}(t_m|\overline{X}_{jk}(t_m), V_{jk})^{\delta_{mjk}} \exp\Big(-\sum_{g\neq j} \Lambda_{jg}(t_m|\overline{X}_{jg}(t_m), V_{jg})\Big) \quad (5)$$

where $\delta_{mjk} = 1$ for a j to k transition and 0 otherwise, $\Lambda_{jk}(t_m|\overline{X}_{jk}(t_m), V_{jk}) = \int_0^{t_m} \lambda_{jk}(s|\overline{X}_{jk}(s), V_{jk})\, ds$, the integrated intensity.

The income of a migrant is only observed while the migrant is employed and depends on the time spend in the country (YSM) and other (possibly varying with time in the host country) covariates

$$\ln W(t) = \theta_0 + \sum_{h=1}^{H} \theta_h YSM_h(t) + \theta_x x(t) + \epsilon(t), \quad (6)$$

where, for a given migrant, the error term is composed of two components, an independently normally distributed idiosyncratic component and a random individual-specific component

$$\epsilon(t) = \eta(t) + v_w.$$

The likelihood contribution from a sequence of income observations over an employment spell is thus

$$L^w\big(W(1), \ldots, W(t)|x(1), \ldots, x(t), v_w\big) = \prod_{s\leq t} \phi\Big(\frac{\ln W(s) - \theta_0 - \sum_{h=1}^{H} \theta_h I_h(s) - \theta_x x(s) - v_w}{\sigma_\eta}\Big) \quad (7)$$

with σ_η being the standard deviation of the idiosyncratic component and $\phi(\cdot)$ the standard normal probability density function.

For the sake of parsimony, we assume that each of the unobserved heterogeneity terms remains the same for recurrent durations of the same type, and we adopt a discrete distribution, i.e. V has the discrete support $(V_1, \ldots, V_M)$ and $p_m = \Pr(V = V_m)$[2]. It is important to note that the V_ms are vectors with $V_m = (V_{12m}, V_{13m}, V_{21m}, V_{23m}, V_{31m}, V_{32m}, V_{wm})'$ including all the possible transitions and the random components of the income equations.

The complete likelihood function for each individual is

$$L = \int L^w(\cdot|V) \prod_{j=e,u,n} \prod_{k \neq j} L_{jk}(V)\, dH_{jk}(V_{jk}) \tag{8}$$

where $H_{jk}(V_{jk})$ is the distribution function of the unobserved heterogeneity.

From the estimates of this correlated competing risk semi-Markov model in Table 5 (first column, and Figure 1) we conclude that years since migration have a positive impact on wages once we have taken into account the correlation between wages and the labor market status. Interestingly, the panel data estimates overestimate the wage growth in the first few years after arrival, but underestimates the impact later on as years of immigration increase. This suggests that selection into and out-of employment plays a prominent role in the early years in the host country. Ignoring this selection process biases the estimates of the wage growth in the host economy. Our next task is account for the out-migration selectivity.

Table 5: Impact of years-since migration (YSM) on log-wage of migrants

	labor-CCRM	full-CCRM
YSM2	-0.141^{**}	-0.065^{**}
	(0.006)	(0.006)
YSM3	0.164^{**}	0.196^{**}
	(0.006)	(0.008)
YSM4	0.333^{**}	0.350^{**}
	(0.007)	(0.009)
YSM5	0.453^{**}	0.452^{**}
	(0.008)	(0.010)
YSM6	0.509^{**}	0.493^{**}
	(0.008)	(0.011)
YSM7	0.562^{**}	0.533^{**}
	(0.010)	(0.013)
YSM8	0.570^{**}	0.579^{**}
	(0.012)	(0.016)
YSM9	0.583^{**}	0.601^{**}
	(0.020)	(0.027)

$^{**}p < 0.01$

[2]To assure that the probability is between zero and one we estimate q_m with $p_m = e^{q_m}/(1 + \sum e^{q_j})$.

4.5 Including selective out-migration

As has previously been highlighted in the literature, out-migration is a major concern for the estimation of wage growth. So we take this into account. Thus, we extend the model from the previous section to include the additional state of living abroad (out-migration) and consider a four-state multistate model:

1. Employed in the host country;
2. Unemployed and receiving benefits in the host country;
3. Out of the labor market (includes both unemployed but not receiving benefits and non labor marker participants) in the host country;
4. Living abroad (Out-migration; i.e. left the host country)

Again these states are mutually exclusive and exhaust all possible destinations. A migrant may leave a state $j = 1, \ldots, 4$ (we ignore repeated immigration) for any of the other destination states. Just as in the previous section we assume that migrant behavior can be described as a semi-Markov competing risk process with individuals moving between the first three states and now with abroad being an absorbing state.[3]

The estimates in Table 5 (second column) and Figure 1 show a positive impact of years since migration on wages once we have taken into account the out-migration selectivity and the correlation between wages and labor market status.[4] However, comparing theses results with the estimates from the multistate model with interdependent labor market transitions without out-migration underscores that selective labor market transitions are as important as accounting for selective out-migration; i.e. this implies that endogenous selection into the labor market is as important as the out-migration endogenous election.

[3]One limitation though is that we assume that unobserved heterogeneity is time invariant, though we model dynamic selection where the probability of leaving and entering employment and the other labor market states changes.

[4]The full tables of estimated coefficients are available from the authors upon request.

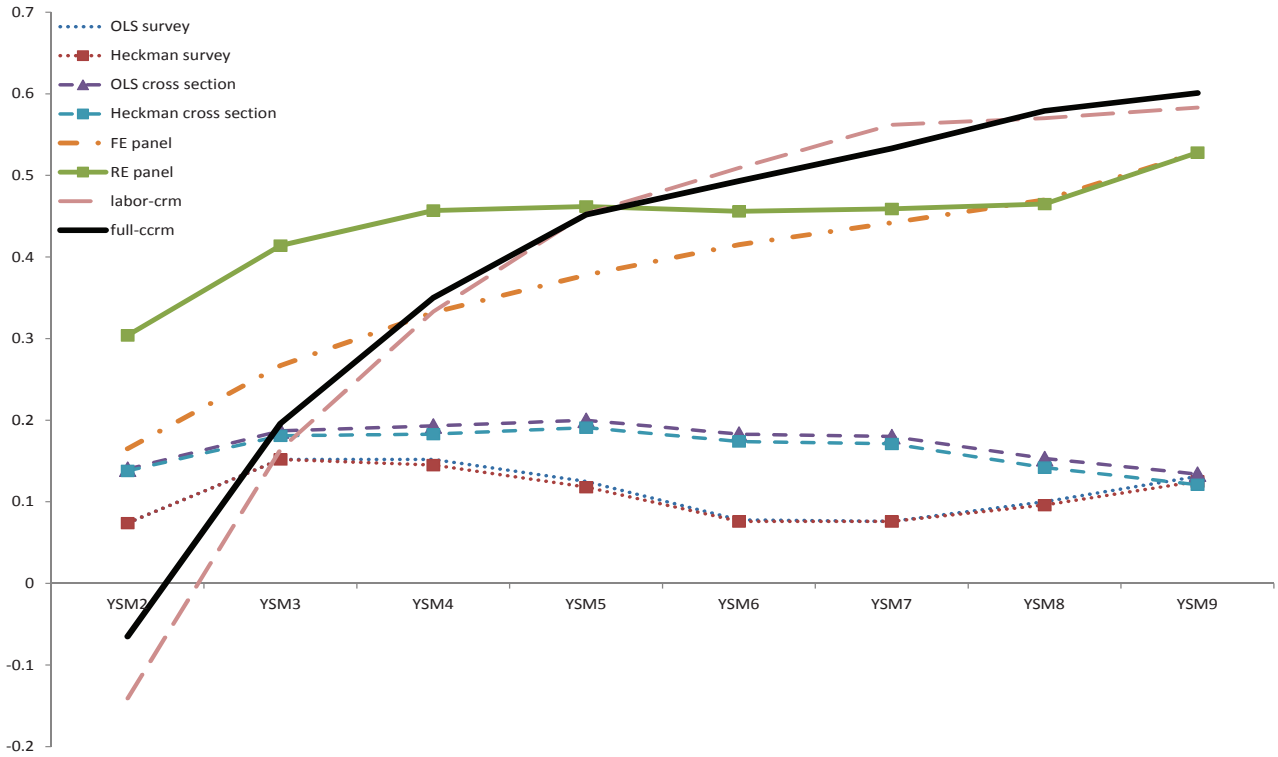

Figure 1: Wage increase by years since migration

5 Distinguishing by migration motive

In order to understand this potential selectivity in out-migration, we differentiate between different types of migrants, who are commonly indistinguishable in the census or commonly used panel data. Tables 6 to 8 provide the characteristics of migrants by type. In Section 5.2 we also distinguish between the region/country of origin given the differences in migration restrictions faced by immigrants; we distinguish between four main groups: immigrants from the EU15 who can move freely in the Dutch labor market, immigrants from the new EU, immigrants from other developed countries (DC) and immigrants from developing countries (LDC). One striking difference, though not surprising, is that almost three quarters of all family migrants are females, compared to only about 30% among labor migrants. Half of the labor migrants are from EU15, and 14% are from the new EU (A8 countries). On the other hand, 34% of family migrants are from LDCs out of which 20% are from Turkey and 18% are from Morocco. Therefore, we additionally categorize the LDC family migrants by these two countries. At the end of the observation window (1/1/2008), 84% of the labor migrants are employed compared to 42% among family migrants, but labor migrants are likely to be more recent (or have fewer YSM).

Not surprisingly only 5% of family migrants were employed at the time of entry but 42% became employed at the end of our observation period. On the other hand, all labor migrants were employed on arrival but only 84% were employed by the end of our observation period. Thus, it is clear that the labor market transitions are different between those two types of migrants and those transition are likely to have different effects on both wage growth and out-migration selection.

Another interesting difference between labor and family migrants is in their income at entry and the change with the YSM. Clearly, the income at entry for family migrants is much lower, even after conditioning for a positive wage. The increase in the average income for family migrants with YSM is due to their increase in paid employment. For neither of the migrant groups we see a clear pattern in the wage with YSM.

Table 6: Descriptive statistics at entry, by migration motive

	Migration motive	
	Labor	Family
EU 15	52.0%	8.3%
new EU	13.7%	6.6%
DC	15.4%	13.4%
LDC	16.4%	34.2%
Turkey	2.0%	19.7%
Morocco	0.5%	17.9%
Female	29.2%	73.8%
Married	24.0%	100%
Age < 25	19.1%	29.4%
Age 25-30	29.6%	25.3%
Age 30-35	21.1%	18.9%
Age 35-40	12.8%	12.0%
Age 40-45	8.1%	6.9%
Age 45-50	5.0%	3.9%
Age 50-65	4.2%	3.6%
Employed	100%	5.1%
Sector[a]		
Agriculture	1.8%	3.9%
Industry	11.5%	6.9%
Construction	1.5%	1.4%
Trade	13.6%	9.7%
Catering	5.1%	4.4%
Transport	5.2%	2.8%
Finance	3.9%	1.8%
Service	43.9%	60.6%
Education	7.7%	4.4%
Health	2.7%	2.2%
Public service	3.1%	2.0%
Total	92,893	101,882

[a] Percentage of all employed.

Table 7: Descriptive statistics at observation end, by migration motive

	Migration motive	
	Labor	Family
YSM		
0–1 year	24.3%	7.3%
1–2 year	18.0%	8.5%
2–3 year	7.6%	8.8%
3–4 year	8.1%	10.6%
4–5 year	7.9%	14.4%
5–6 year	9.1%	13.7%
6–7 year	9.4%	13.4%
7–8 year	8.7%	12.1%
8–9 year	6.8%	11.2%
Employed	83.6%	41.8%
Sector[a]		
Agriculture	1.6%	1.9%
Industry	11.7%	10.5%
Construction	1.7%	2.7%
Trade	14.8%	13.2%
Catering	3.8%	7.2%
Transport	6.0%	3.6%
Finance	5.1%	1.7%
Service	39.5%	46.1%
Education	7.8%	2.9%
Health	4.6%	7.3%
Public service	3.3%	3.1%
Total	48,599	80,801

[a] Percentage of all employed at end.

Table 8: Descriptive statistics income

	Average		Average (>0)[a]		Average (>0, Emp)[b]	
	labor	Family	labor	Family	labor	Family
entry	€2276	€ 98	€3331	€1572	€3331	€1835
last observation	€2572	€ 771	€4435	€1645	€4527	€1887
observation end	€3814	€ 890	€4456	€1614	€4524	€1846
YSM						
0–1 year	€3670	€ 383	€3970	€1781	€3981	€1889
1–2 year	€3784	€ 616	€4419	€1580	€4451	€1771
2–3 year	€3955	€ 904	€4702	€1943	€4760	€2240
3–4 year	€4453	€ 851	€5331	€1572	€5427	€1811
4–5 year	€4574	€ 885	€5613	€1519	€5754	€1730
5–6 year	€3395	€ 916	€4189	€1522	€4308	€1739
6–7 year	€3569	€1082	€4311	€1710	€4451	€1978
7–8 year	€3549	€1034	€4291	€1589	€4337	€1812
8–9 year	€3844	€1044	€4591	€1571	€4798	€1805

[a] Only for migrants with positive income.
[b] Only for employed migrants with positive income.

5.1 Findings by migration motive

Table 9 shows the estimates for three models: the fixed effects model, the labor-CCRM and the full-CCRM model by different migrant types. The results for all the considered models for the two migration motives are depicted in Figure 2 and 3. First, it should be emphasized that the estimates for the whole sample (both types of immigrants) shown in Figure 1 are not a simple weighted average of the labor and family migrants (except for the OLS results). The selection processes for both types of immigrants may be different and thus when we combine them (into the whole sample) we might lose part of the selection because we assume it similar for all immigrants. Secondly, for both types of immigrants, the fixed effect panel data estimates are the closest to the CCRM estimates but are still biased. For labor immigrants the FE model underestimates the wage growth suggesting positive out-migration selection whilst the opposite is true for family immigrants indicating negative selection of out-migration amongst family immigrants. Both CCRM models produce similar estimates but the labor-CCRM overestimates slightly wages as years-since-migrating increase. Finally, except for the wage grows in the first three years the wages of family migrants growth faster than the wages of labor migrants. Of course, many family migrants are not employed at entry and therefore have no wages to begin with and as a result experience a greater opportunity for wages to grow. However, ignoring the interdependence of the labor market processes leads to mis-estimation of the wage growth. Overall, our estimates underscore the importance of differentiating between different types of immigrants given that changes in type lead to different labor supplies and out-migration probabilities impacting on earnings profiles.

Table 9: labor vs Family migrants

	Labor			Family		
	FE	labor-CCRM	full-CCRM	FE	labor-CCRM	full-CCRM
YSM2	0.098**	0.191**	0.186**	0.333**	−0.367**	−0.379**
	(0.001)	(0.006)	(0.007)	(0.002)	(0.008)	(0.011)
YSM3	0.161**	0.260**	0.260**	0.487**	0.070**	0.048**
	(0.001)	(0.007)	(0.008)	(0.003)	(0.009)	(0.012)
YSM4	0.215**	0.319**	0.310**	0.572**	0.294**	0.275**
	(0.002)	(0.008)	(0.010)	(0.004)	(0.010)	(0.013)
YSM5	0.268**	0.381**	0.373**	0.623**	0.446**	0.420**
	(0.002)	(0.009)	(0.012)	(0.005)	(0.011)	(0.015)
YSM6	0.321**	0.430**	0.417**	0.651**	0.534**	0.491**
	(0.002)	(0.010)	(0.013)	(0.006)	(0.012)	(0.016)
YSM7	0.370**	0.466**	0.415**	0.668**	0.595**	0.547**
	(0.003)	(0.012)	(0.016)	(0.007)	(0.013)	(0.018)
YSM8	0.410**	0.480**	0.409**	0.693**	0.629**	0.550**
	(0.003)	(0.015)	(0.021)	(0.008)	(0.016)	(0.022)
YSM9	0.493**	0.492**	0.469**	0.766**	0.624**	0.567**
	(0.007)	(0.026)	(0.037)	(0.011)	(0.025)	(0.018)

$^{+}p < 0.05$ and $^{**}p < 0.01$

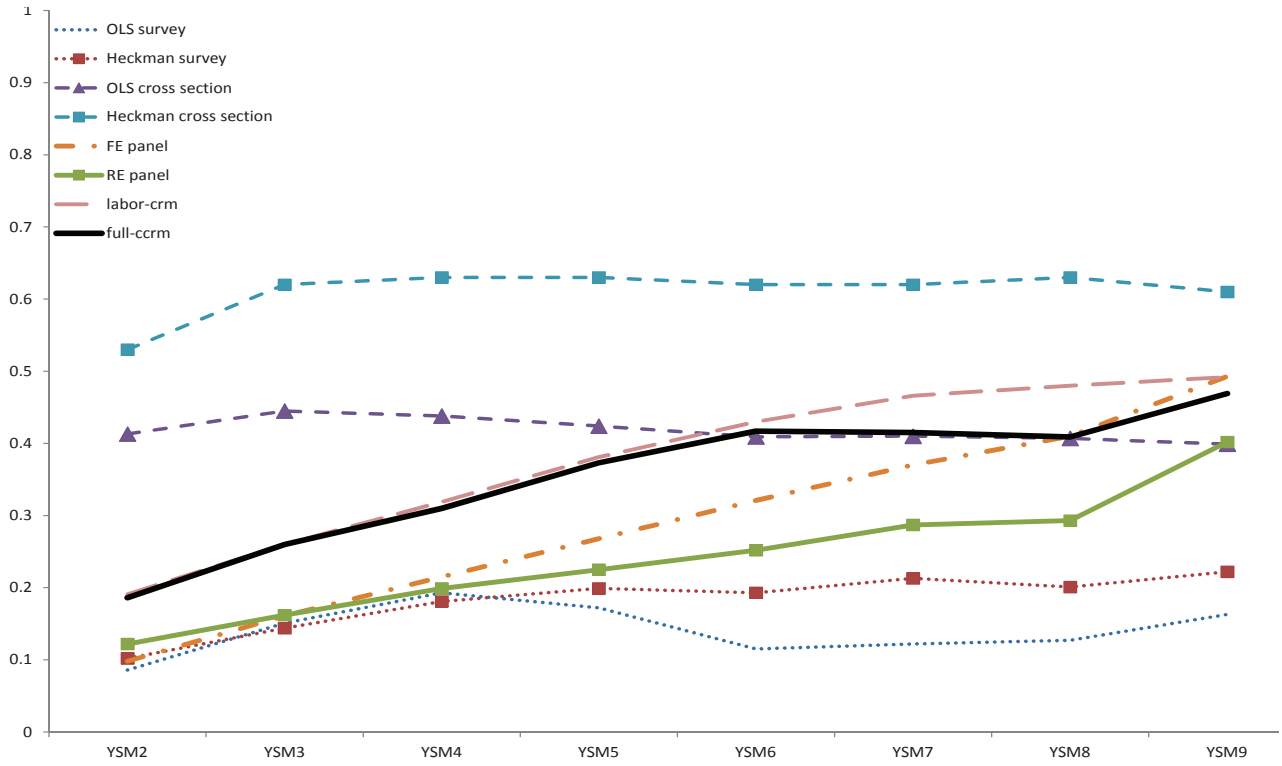

Figure 2: Wage increase by years since migration, labor migrants

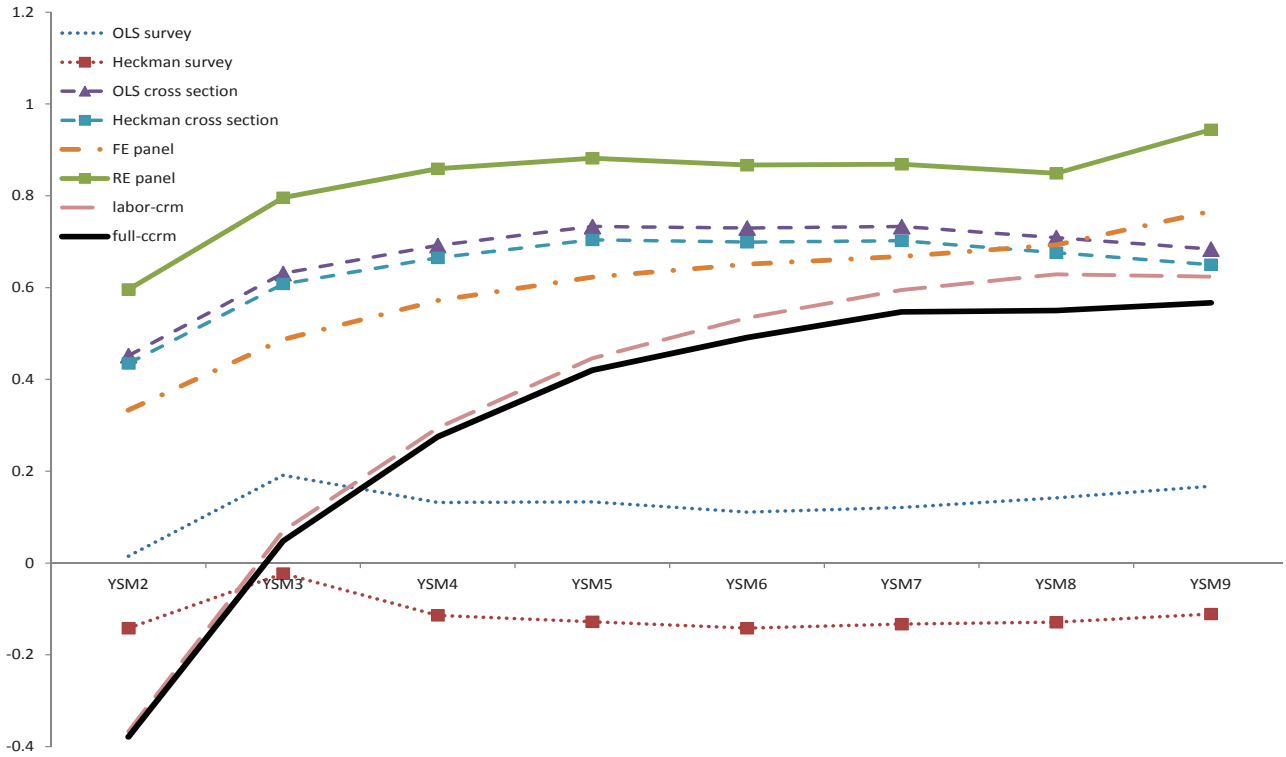

Figure 3: Wage increase by years since migration, family migrants

5.2 Findings by region of origin and migration motive

Previous evidence shows considerable differences in the relationship between earnings and out-migration by country of origin. Interestingly, when we distinguish between labor migrants by region/country of origin, the full-CCRM estimates (Table 10) highlight that labor migrants from the EU15 experience the smallest wage growth compared to those from the new EU, DCs and LDCs. This is consistent with Bijwaard et al (2014) who find that EU15 labor immigrants in the Netherlands tend to have the highest unemployment incidence. Table 11 shows that the returns to YSM for family migrants increase over time and tend to be higher than those of labor migrants. Indeed, immigrants from the same country/region, experience different wage growth depending on whether they are labor or family migrants; i.e. the region/country of origin matters as much as the migrant type for the wage growth. This confirms that out-migration selectivity is host-origin country specific.

Table 10: Full-CCRM: Labor migrants

	EU 15	new EU	DC	LDC
YSM2	0.151**	0.157**	0.226**	0.235**
	(0.004)	(0.009)	(0.008)	(0.008)
YSM3	0.182**	0.255**	0.274**	0.283**
	(0.005)	(0.012)	(0.010)	(0.010)
YSM4	0.194**	0.323**	0.336**	0.376**
	(0.006)	(0.014)	(0.012)	(0.011)
YSM5	0.210**	0.446**	0.402**	0.473**
	(0.007)	(0.017)	(0.014)	(0.013)
YSM6	0.216**	0.534**	0.509**	0.548**
	(0.008)	(0.020)	(0.016)	(0.014)
YSM7	0.208**	0.628**	0.592**	0.622**
	(0.009)	(0.024)	(0.019)	(0.017)
YSM8	0.167**	0.646**	0.605**	0.660**
	(0.012)	(0.035)	(0.027)	(0.023)
YSM9	0.144**	0.726**	0.652**	0.660**
	(0.021)	(0.073)	(0.051)	(0.044)

$^{+}p < 0.05$ and $^{**}p < 0.01$

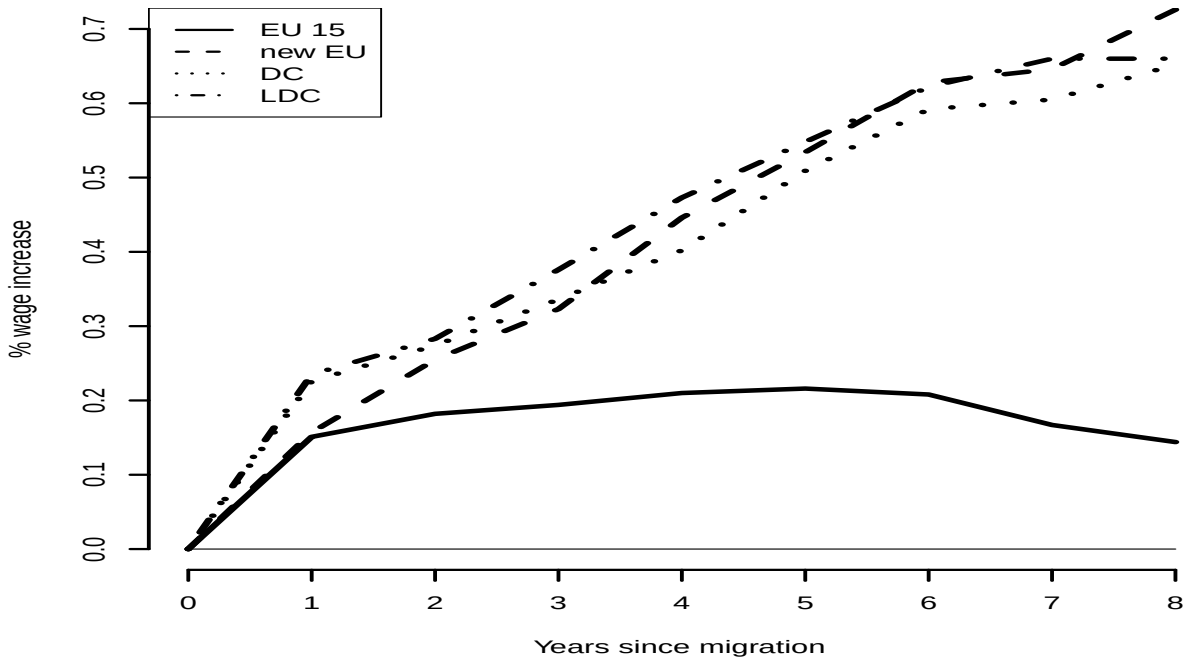

Figure 4: Wage increase by years since migration, labor migrants by region

Table 11: Full-CCRM: Family migrants

	EU 15	new EU	DC	LDC	Turkey	Morocco
YSM2	-0.501^{**}	-0.556^{**}	-0.676^{**}	-0.542^{**}	-0.245^{**}	-0.087^{**}
	(0.022)	(0.022)	(0.023)	(0.011)	(0.010)	(0.010)
YSM3	-0.123^{**}	-0.145^{**}	-0.167^{**}	-0.071^{**}	0.147^{**}	0.287^{**}
	(0.026)	(0.025)	(0.025)	(0.012)	(0.011)	(0.011)
YSM4	-0.017	0.116^{**}	0.099^{**}	0.194^{**}	0.313^{**}	0.498^{**}
	(0.030)	(0.029)	(0.028)	(0.013)	(0.012)	(0.012)
YSM5	0.116^{**}	0.218^{**}	0.267^{**}	0.360^{**}	0.450^{**}	0.624^{**}
	(0.034)	(0.034)	(0.030)	(0.015)	(0.013)	(0.014)
YSM6	0.243^{**}	0.347^{**}	0.328^{**}	0.465^{**}	0.530^{**}	0.690^{**}
	(0.037)	(0.037)	(0.032)	(0.016)	(0.014)	(0.015)
YSM7	0.294^{**}	0.386^{**}	0.439^{**}	0.533^{**}	0.594^{**}	0.739^{**}
	(0.041)	(0.043)	(0.036)	(0.018)	(0.016)	(0.017)
YSM8	0.291^{**}	0.434^{**}	0.501^{**}	0.584^{**}	0.619^{**}	0.780^{**}
	(0.051)	(0.056)	(0.044)	(0.021)	(0.019)	(0.020)
YSM9	0.292^{**}	0.467^{**}	0.502^{**}	0.626^{**}	0.638^{**}	0.827^{**}
	(0.075)	(0.091)	(0.070)	(0.034)	(0.032)	(0.032)

$^{+}p < 0.05$ and $^{**}p < 0.01$

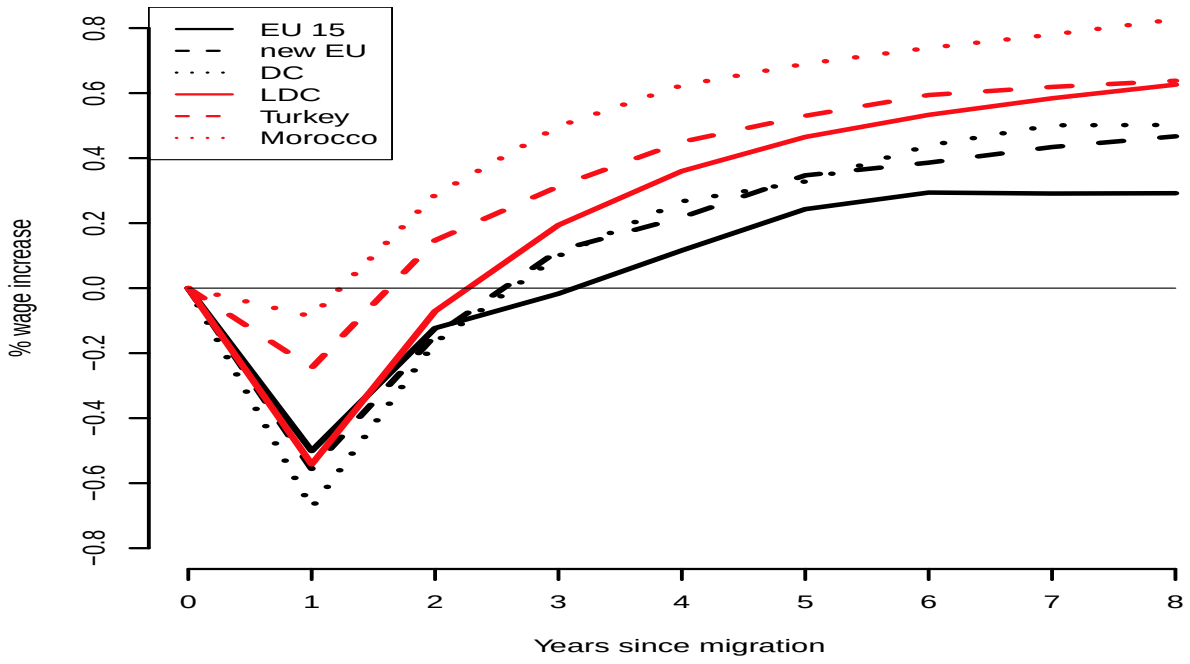

Figure 5: Wage increase by years since migration, family migrants by region

5.3 Distinguishing by Initial income

Finally, an important issue that has arisen in the literature is whether various income groups experience different income trajectories. In particular, whether immigrants who have low initial earnings suffer from persistent labor-market disadvantages or experience faster wage growth. For example, Borjas (1999) and Hu (2000) find that immigrants who have lower initial earnings quickly make up a large part of the deficit relative to their immigrant counterparts. In Table 12 and Figure 6 we show the full-CCRM estimates for three income groups where income refers to that at time of entry. It is clear that all the groups experience wage growth but, not surprisingly, the lowest income group experiences the biggest wage growth as previously documented in the literature, this is the case even after controlling for the endogeneity of out-migration selection and the correlated labor market history.

Table 12: CCRM: labor migrants by initial income

	Low income	Medium income	High income
YSM2	0.594**	0.020**	0.009
	(0.008)	(0.004)	(0.005)
YSM3	0.710**	0.105**	0.059**
	(0.010)	(0.005)	(0.006)
YSM4	0.802**	0.170**	0.111**
	(0.012)	(0.005)	(0.007)
YSM5	0.872**	0.255**	0.193**
	(0.013)	(0.006)	(0.009)
YSM6	0.935**	0.329**	0.275**
	(0.015)	(0.007)	(0.010)
YSM7	0.998**	0.386**	0.334**
	(0.017)	(0.008)	(0.013)
YSM8	0.986**	0.421**	0.366**
	(0.023)	(0.010)	(0.018)
YSM9	0.961**	0.502**	0.430**
	(0.041)	(0.019)	(0.035)

Low income (Monthly income) $<$ €1000; Medium income €1000- €3000; High income $>$ 3000 $^{+}p < 0.05$ and $^{**}p < 0.01$

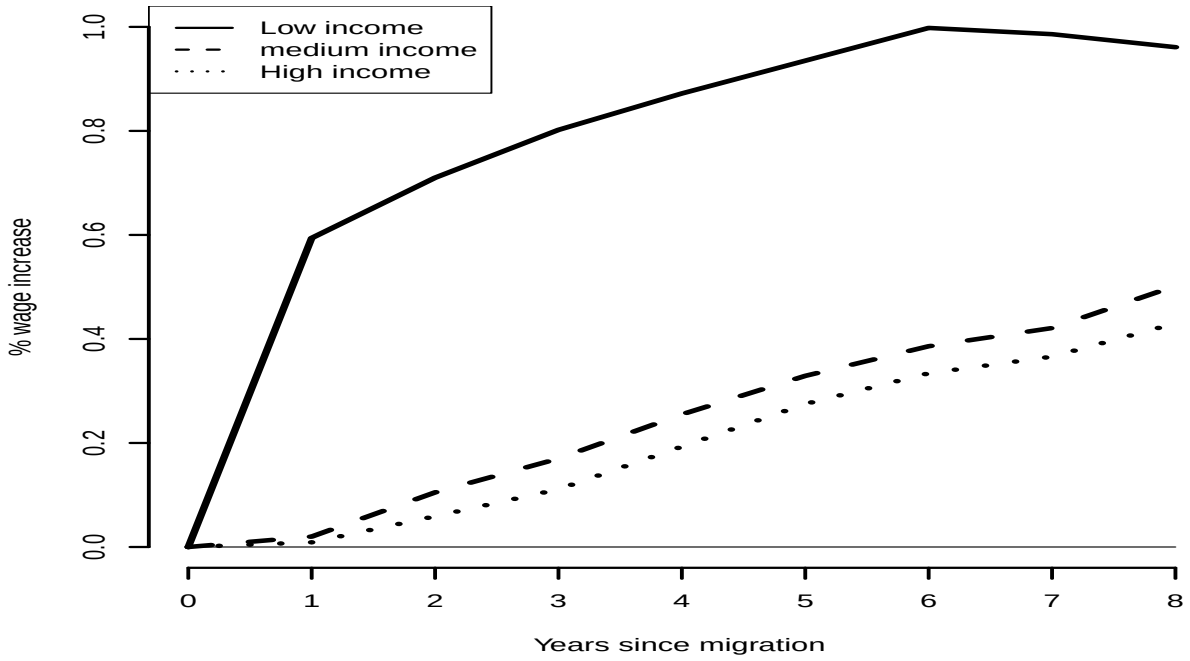

Figure 6: Wage increase by years since migration, labor migrants by initial income at entry

6 Conclusion

The performance of immigrants in the labor market, in particular their earnings' growth, is an important indicator of their contribution to the economy. However, examining the earnings profiles of immigrants over their migration cycle is challenging. As acknowledged in the previous literature, the non random out-migration leads to a selection bias that is hardly addressed due to a lack of data. This paper examines immigrant wage growth taking into account the out-migration selection using administrative data from the Netherlands. More importantly, we also address an additional challenge that has been overlooked in this literature namely, that out-migration might be correlated with the labor supply and earnings of immigrants. We take into account the correlation between three important factors which has not been considered before: the labor market status of the migrant, the wages earned and the out-migration decision as they can all impact on the estimated wage growth of current immigrants. Finally, we distinguish between labor and family migrants whose labor market and out-migration behaviors are different.

We follow the standard literature and treat our data as a single cross section, a repeated cross section and a panel before using the administrative longitudinal data of the full immigrant population

to estimate a system of correlated competing risk (CCRM) models. The findings suggest that the fixed effects (FE) panel data are closest to our CCRM estimates. However, the FE model over-estimates the wage growth in the early years in the host country indicating negative selection of out-migration early on in the migration cycle, but under-estimates the wage growth in longer migration durations suggesting positive selection of out-migration taking place in longer durations. This is consistent with a story where unsuccessful immigrants leave in the first 2-3 years after arrival whilst the more successful immigrants leave later on after having met their target savings or human capital acquisition. Distinguishing between types of migrants sheds further light: the FE panel model under-estimates the wage growth of labor migrants but over-estimates the wage growth of family migrants reinforcing the dichotomy between the various immigrant groups where labor out-migrants are positively selected but family out-migrants are negatively selected. Finally, our findings also underscore that if the endogeneity of the labor supply, wages, and out-migration were ignored then estimating immigrant wage growth would be problematic and would lead to biased estimates. Indeed the endogeneity of labor supply is as important as the out-migration selection for wage estimations.

References

Abramitzky, R., L. P. Boustan, and K. Eriksson (2012). A nation of immigrants: Assimilation and economic outcomes in the age of mass migration. Working paper 18011, NBER.

Biavaschi, C. (2012). Recovering the counterfactual wage distribution with selective return migration. Discussion paper 6795, IZA.

Bijwaard, G. E. (2010). Immigrant migration dynamics model for The Netherlands. *Journal of Population Economics 23*, 1213–1247.

Bijwaard, G. E., C. Schluter, and J. Wahba (2014). The impact of labour market dynamics on the return–migration of immigrants. *Review of Economics and Statistics 96*, 483–494.

Bijwaard, G. E. and J. Wahba (2014). Do high-income or low-income immigrants leave faster? *Journal of Development Economics* 108, 54-68.

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

Borjas, G. J. (1989). Immigrant and emigrant earnings: A longitudinal study. *Economic Inquiry 27*, 21–37.

Borjas, G. J. (1999). The economic analysis of immigration. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics, Volume 3A*, Chapter 28. Amsterdam: North–Holland.

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

Dustmann, C. (1995). Return migration: The European experience. *Economic Policy 22*, 214–250.

Dustmann, C. and J.-S. Görlach (2014). Selective outmigration and the estimation of immigrants earnings profiles. In B. R. Chiswick and P. W. Miller (Eds.), *Handbook on the Economics of International Migration, 1A*, Chapter 10, pp. (forthcoming). Amsterdam: North–Holland.

Edin, P.-A., R. J. Lalonde, and O. Åslund (2000). Emigration of immigrants and measures of immigrant assimilation: Evidence from Sweden. *Swedish Economic Review 7*, 163–204.

Hu, W. Y. (2000). Immigrant earnings assimilation: Estimates from longitudinal data. *American Economic Review 90*, 368–372.

Jasso, G. and M. R. Rosenzweig (1982). Estimating the emigration rates of legal immigrants using administrative and survey data: The 1971 cohort of immigrants to the US. *Demography 19*, 279–290.

Jensen, P. and P. J. Pedersen (2007). To stay or not to stay? out–migration of immigrants from Denmark. *International Migration 45*, 87–113.

LaLonde, R. J. and R. H. Topel. The assimilation of immigrants in the U. S. labor market. In G. J. Borjas and R. B. Freeman (Eds.), *Immigration and the Workforce: Economic Consequences for the United States and source areas*. Chicago: University of Chicago Press.

Lubotsky, D. (2007). Chutes or ladders? a longitudinal analysis of immigrant earnings. *Journal of Political Economy 115*, 820–867.

Massey, D. S., J. Arango, G. Hugo, A. Kouaouci, A. Pellegrino, and J. E. Taylor (1993). Theories of international migration: A review and appraisal. *Population and Development Review 19*, 431–466.

Pischke, J. S. (1992). Assimilation and the earnings of guest workers in Germany. Discussion paper 92-17, ZEW, Mannheim.