DISCUSSION PAPER SERIES

Do High-Income or Low-Income Immigrants Leave Faster?

Govert E. Bijwaard
Jackline Wahba

November 2013

I Z A

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Do High-Income or Low-Income Immigrants Leave Faster?

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

Jackline Wahba
University of Southampton
and IZA

ABSTRACT

Do High-Income or Low-Income Immigrants Leave Faster?[*]

We estimate the impact of the income earned in the host country on return migration of labor migrants from developing countries. We use a three-state correlated competing risks model to account for the strong dependence of labor market status and the income earned. Our analysis is based on administrative panel data of recent labor immigrants from developing countries to the Netherlands. The empirical results show that intensities of return migration are U-shaped with respect to migrants' income, implying a higher intensity in low- and high-income groups. Indeed, the lowest-income group has the highest probability of return. We also find that ignoring the interdependence of labor market status and the income earned leads to an overestimating the income effect on departure.

JEL Classification: F22, J61, C41

Keywords: migration dynamics, labour market transitions, competing risks,
              immigrant assimilation

Corresponding author:

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

1 Introduction

The literature on international migration is large and growing, but only recently attention has been paid to temporary migration as many migrants, in fact, return. The limited theoretical literature on return migration provides several explanations for why migrants return. On the one hand, return migration is seen as planned and part of optimal decision making to maximize total utility over the whole life cycle where return migration is motivated by locational preference for home country, e.g. consumption, or differences in relative prices in host and home country (e.g. Galor and Stark (1991) and Dustmann (1997)). Thus, migrants migrate temporarily to accumulate resources, or skills, for later use in the home country. On the other hand, another strand of this emerging literature sees return migration as unplanned and the result of failure either due to imperfect information about the host country in terms of labor market prospects or the cost of living, or the inability to fulfil the migration plans in terms of target savings (see Borjas and Bratsberg (1996)).

An interesting issue that has been understudied is the relationship between the migration duration and migrant's income abroad. Although there is a consensus that migration is driven by the wage differential between the host and the home country, the effect of wages (or income) on return migration is ambiguous. Migrants would, on the one hand, like to extend their stay overseas as a response to higher wages; on the other hand, the gain from staying longer abroad decreases. As a consequence, higher wages abroad may have a positive or a negative effect on migration duration.

This paper contributes to this literature by using unique data that circumvent several data problems encountered in previous studies. 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 and the exact detailed information on their labor market status and income. This enables us to address our question of interest on the effect of income on migration duration in a novel way that takes into account the changing nature of income experienced by migrants, and control for the correlation between the potential endogenous labor market status of the migrant and the return decision. Given the substantial heterogeneity between immigrants from different origins and the corresponding variation in the immigration policies that impact on the free movement of immigrants and hence their return migration, we limit our focus to labor immigrants from less developed countries (LDC).

The empirical evidence on the effect of income on migrant's duration abroad is rather limited due to lack of data and is mixed. For example, Borjas (1989) finds among the foreign-born in the United

States that higher earnings are associated with less return migration. By contrast, Dustmann (2003) shows that immigrants in Germany return earlier when the wage level in the host country increases, whilst Nekby (2006) finds a U relation between income and out-migration. However, Constant and Massey (2003) find no statistically significant relationship between earnings and migrant returns in Germany, although migrants who are unemployed are more likely to return. Furthermore, Gibson and McKenzie (2011) who successfully tracked down a high proportion of the very top performers in secondary school from 1976 to 2004 from three Pacific countries, find that narrow measures of income gains play a very minor role in determining which of the highly skilled return.

These previous papers relied on longitudinal data that suffer from high attrition rates and lack information on the exact timing of the migration moves and only reveal whether the migrant is still in the country at the interview date. We use data from Statistics Netherlands, which includes data on a monthly basis, the labor market status and income of the migrants. The timing of both labor market status changes and migration status changes allows us to construct the full labor market and migration history. The duration in each labor market state forms the basis of our analysis. Duration, or event history, models have been used extensively for demographic analysis but are rather limited in migration studies and analysis of return migration is even scarcer.

In this paper we investigate whether it is the high-income or low-income migrants who leave faster. We examine the extent to which the length of migration stay of migrants differs with regard to their initial income level in the host country. To account for the strong dependence between labor market status and income earned, we distinguish between three labor market statuses: employed, unemployed and non-participation and estimate a three-state correlated competing risks model. We control for unobserved correlated heterogeneity in the labor market and income earned at the host country, and migration processes. Given the diversity of immigrants' background, we limit the analysis to labor immigrants from developing countries since the behavior of those immigrants is paramount for policymakers. We also control for home country circumstances by using time varying GDP per capita and economic growth.

This analysis has a number of interesting and important implications for migration policies. Who leaves faster? Is it the 'successful' migrants or the 'unsuccessful' ones? Our empirical results show that return intensities are U-shaped with respect to income, implying a higher intensity in low- and high- income groups. Indeed, the findings suggest that the low-income group has the highest intensity of return. This U shape is found at different migration durations, although the intensities of return decline after 5 - 6 years in the Netherlands. Interestingly, our simulations comparing immigrants from

the main five less developed countries of origin (India, China, Turkey, South Africa and Morocco) find consistent evidence of this U-shaped relationship between initial income and return, with the lowest-income group having the highest intensity, followed by the high income groups. This is consistent with having successful high-income migrants leaving once they have earned their savings or human capital accumulation targets, whilst at the same time, the low-income migrants returning as a result of their limited success. These findings provide evidence of brain circulation as we find high earners having shorter migration duration, and also finding low-income immigrants leaving quickly dampens the concern by many about the fiscal burden of low income immigrants. Finally, our results highlight that ignoring the interdependence of labor market status and incomes earned leads to underestimation of the impact of low income and overestimation of the impact of high income.

The outline of the paper is as follows. We consider briefly the related literature and conceptual framework next. In Section 3, we present the data and discuss briefly recent migration to the Netherlands. In Section 4 we present the results of estimating a simple standard duration model that ignores the possible endogeneity of the labor market status and the income earned. Section 5 spells out the correlated competing risks model (CCRM) which takes this endogeneity into account. Section 6 considers the comparison of important labor market indicators by income status using microsimulation based on the estimated CCRM. Section 7 focusses on the implications for the main countries of origin. The last section concludes.

2 Related Literature and Conceptual framework

Much of the economic research considers migration as permanent (see e.g. Chiswick 1978, Massey et al. 1993 and Borjas 1999). Nevertheless, the level of return migration has been high both in the US and Europe. Jasso and Rosenzweig (1982) report that of the 1971 cohort of immigrants to the US, almost fifty percent returned by 1979. Dustmann (1995) has demonstrated the relevance of return migration in the European context. In the Netherlands, recent migrants also show a high return rate (see Bijwaard (2010)).

Several competing theories have been advanced to explain the impact of the income level of migrants in the host country on their propensity to return. According to one strand of literature, return migration is planned and part of an optimal strategy to maximize life-time utility characterized by a preference for source country consumption (see e.g. Galor and Stark (1991), Dustmann (1997), Dustmann (2003), and Dustmann and Weiss (2007)). Return migration by target savers is but one example. Thus, migrants are viewed as target earners who return home after their target is reached

and hence high income migrants would return faster. A fundamentally different mechanism is based on mistaken expectations about, and immediate failure on the host country's labor market, leading to an 'unplanned' return (Borjas and Bratsberg (1996)). According to this view, return migrants are "failures" and low income migrants are more prone to return faster.

Empirical work focusing on the effect of migrant income on the return decision is rather limited. Borjas (1989) using longitudinal data from the 1972 - 1978 Survey of Natural and Social Scientists and Engineers, finds among the foreign-born in the US that higher earnings are associated with less return migration. Yang (2006) too finds similar qualitative results exploiting a unique quasi-experiment to distinguish between these potential explanations for return migration. He examines how the return decisions of the Philippine migrants respond to major and unexpected exchange rate shocks (due to the 1997 Asian financial crisis). He finds that more favorable exchange rate shocks which can be interpreted as higher income lead to fewer migrant returns. Contrary to that, Dustmann (2003) analyzes optimal migration durations in a model, which rationalizes the decision of the migrant to return to his home country, despite persistently higher wages in the host country. He shows that, if migrations are temporary, the optimal migration duration may decrease if the wages increase based on a panel of immigrants to Germany over a 14-year period. Nekby (2006), using data on registered emigration from Sweden from 1991-2000, finds that although emigrants in general have higher adjusted mean income levels, up to the age of 35-40, than non-emigrants, onward migrants have lower predicted income levels across the age distribution due to this groups relatively low employment levels in Sweden.

However, a few studies find no income effect on return migration. Constant and Massey (2003) find no statistically significant relationship between earnings and migrant returns in Germany, although migrants who are unemployed are more likely to return. Gibson and McKenzie (2011) successfully tracked down a high proportion of the very top performers in secondary school over 1976 to 2004 from three Pacific countries. The results reveal for both the initial decision to emigrate and the decision to return, income gains play a very minor role in determining which of the highly skilled migrate and return, whereas preference variables are strong predictors. None of those previous studies adopt a structural approach and use competing risks model.

Closer to our interest is Bijwaard (2009) and Bijwaard et al. (2013). Bijwaard (2009) considers the correlation between migration decisions and labor market status transitions. Bijwaard et al. (2013) estimate the causal effect of unemployment on the return decision in the Netherlands. Neither of the studies examines the effect of migrants' income on migration duration and the intensity of return.

3 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 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 register 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 in other administrative registers we obtain a panel.

In addition to the date of entry and exit, the administration 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 among the following motives: labor-migrants, family migrants, student immigrants, asylum seekers (and refugees), and immigrants for other reasons. See Bijwaard (2010) for an extensive descriptive analysis of the various migration motives. In particular, about 23% of all non-Dutch immigrants in the age group 18 - 64 are labor migrants. Given our interest in the effect of migrant income on return, we focus exclusively on labor migrants and restrict our analysis to those immigrants who are employed in the Netherlands within three months of their entry. Non-labor migrants have different motives for migrating, such as family or study. Hence the effect of income on their return is different. Furthermore, given the substantial heterogeneity between immigrants from different origins and the corresponding variation in the immigration policies that impact on the free movement of immigrants and hence their return migration, we limit our focus to labor immigrants from less developed countries (LDC).

Although, in principle, the exact date of emigration is known, some migrants do not officially inform the authorities when they leave. The departure of these non-complying individuals is registered as an 'administrative removal' after the authorities have assessed that the migrant has left the municipality without showing up in the files of another municipality in The Netherlands or as an emigrant. These administrative removals are included among emigration and they add up to about 38% of all emigrations and 73% of these administrative removed migrants have no observed income in the country. We conjecture that the majority of these migrants have left the country shortly after they stopped receiving income (either earnings or benefits). For those who still have income until they are administratively removed we assume that they left at that exact date. For those who are both

administrative removed and have "zero income at last observed time", we assume that the migrant has left *before* the date the administrative removal is recorded, and *after* the last date of any observed change in the observed characteristics (e.g. labor market status, housing and marital status). Such limited information is equivalent to *interval-censored* data. For interval-censored data the exact end of duration is unknown, but it is known that the duration ended in some time period. We have explicitly addressed the issue of administrative removals in the formulation of the likelihoods below.

The immigration register is linked by Statistics Netherlands to the Municipal Register of Population (Gemeentelijke Basisadministratie, GBA) and to their 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 information (on a monthly basis) on the labor market status, income, industry sector, housing and household situation. To capture, country of origin's economic situation, we use annual GDP per capita and GDP growth rate by country of origin from the World Bank, World Development Indicators. To control for the host country's labor market, national unemployment rates are used. We also control for the potential immigrant cohort effects, by using the unemployment rate in the Netherlands at the time of immigration. We distinguish three labor market categories: (1) employed and self-employed, (2) unemployed but receiving benefits and (3) non-participating (which includes those unemployed who are illegible for any benefits and those with no income). Note that LDC immigrants entering during our observation window do not qualify for social benefits straight away, as eligibility requires sufficiently long employment or residence durations.

3.1 Descriptive statistics

Table 1: Descriptive statistics at entry, LDC labor migrants

| | Initial income: from low to high | | | | | | |
	Inc 1	Inc 2	Inc 3	Inc 4	Inc 5	Inc 6	Inc 7
Female	28.1%	22.5%	23.3%	21.5%	19.2%	19.9%	15.5%
Single	79.4%	76.2%	72.2%	67.7%	65.8%	53.9%	45.9%
Married	18.9%	22.5%	26.5%	31.9%	33.5%	43.4%	53.1%
av. age	30.3	30.4	31.0	31.6	32.5	34.8	36.9
GDP pc.	$2976	$3544	$4051	$4406	$4791	$5943	$5483
GDP growth	5.2%	5.8%	6.0%	6.2%	6.0%	5.2%	5.2%
Distribution	32.3%	23.0%	17.5%	9.8%	5.1%	3.0%	9.4%

Inc 1: Monthly income < €1000; Inc 2: Monthly income €1000-€2000; Inc 3: Monthly income €2000-€3000; Inc 4: Monthly income €3000-€4000; Inc 5: Monthly income €4000-€5000; Inc 6: Monthly income €5000-€6000; Inc 7: Monthly income > €6000;

First, we provide an overview of our data. Table 1 shows various migrant characteristics by initial

income group of our sample of 16,974 labor immigrants from LDCs. Almost 77% are men and they are most often single (71%). The immigrants are relatively young, with 16% younger than 25 and 48% younger than 30. The main countries of origin of our LDC labor immigrants are: India (19%), China (10%), South Africa (8%), Brazil (4%), Taiwan (4%) and Morocco (3%). The average income of the migrants at the time of arrival is €2751, with 32% earning €1000 or less monthly and another 23% earning only between €1000 and €2000 a month. The average GDP per capita in the home country is $3151 and the average growth rate of the country of origin is 4.8%. Interestingly, the proportion of women is the highest in the lowest-income group. Moreover, low earners are more likely to be single and younger compared with the high earners. Indeed, there seems to be a correlation between the GDP per capita of a country of origin and the migrant income group.

The unconditional distribution of the immigration duration (Figure 1) depicts the Kaplan-Meier estimates of the survival probabilities by income immigrant group. All groups look similar for durations. However, for the top earners (more than €6000), they show the highest survival rate up to 24 months, then at longer durations they have the lowest staying incidence. The bottom income immigrant group tends to have the highest exit rate in the first 2-3 years but later they become the least likely to leave.

Figure 2 shows the Kaplan-Meier estimates of the survival probabilities, from employment, and the cumulative incidence functions by labor market status and income immigrant group. Those estimates show that the survival in employment is the lowest over time for the lowest earner groups. They move more often to non-participation and unemployment. But migrants with a higher initial income leave the country sooner. However, those figures do not take into account the correlation between the labor market status, the change in earned income and survival.
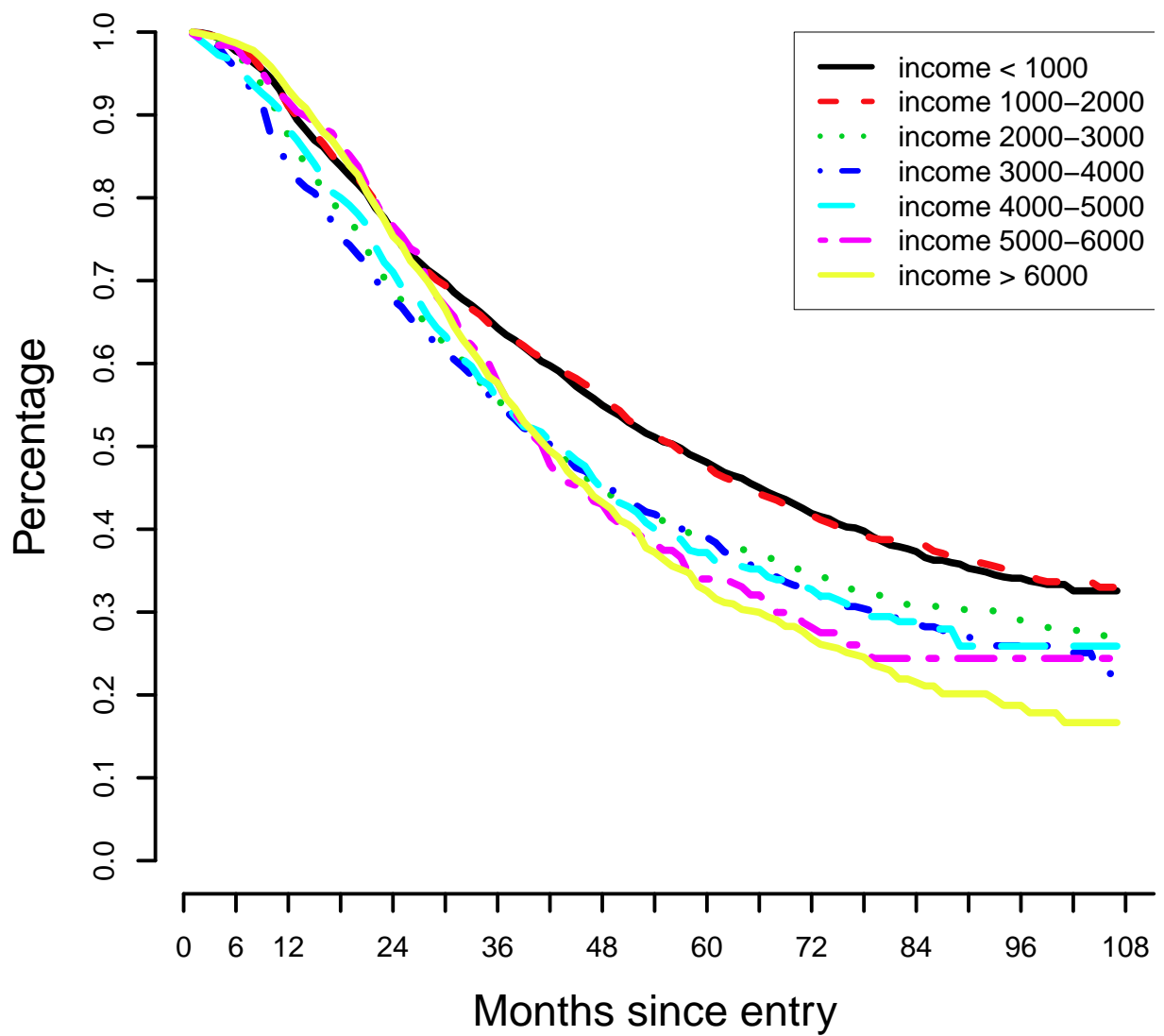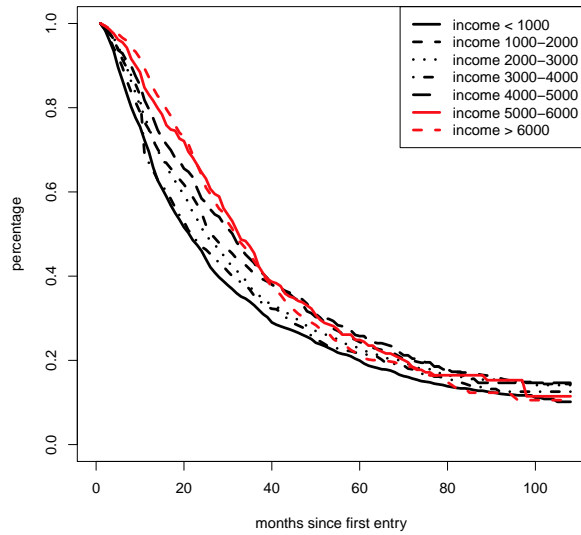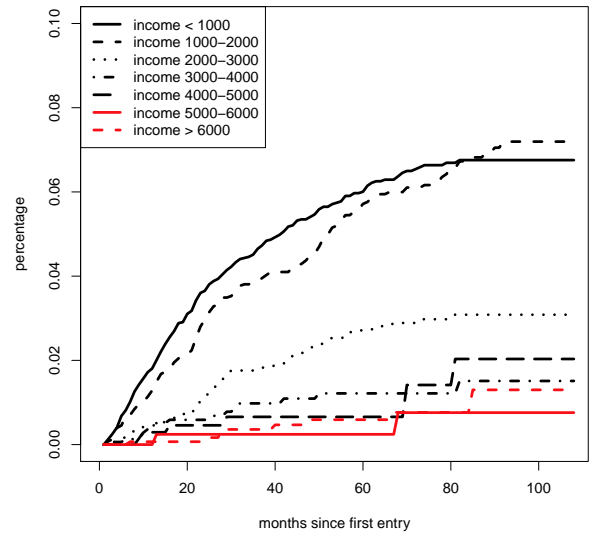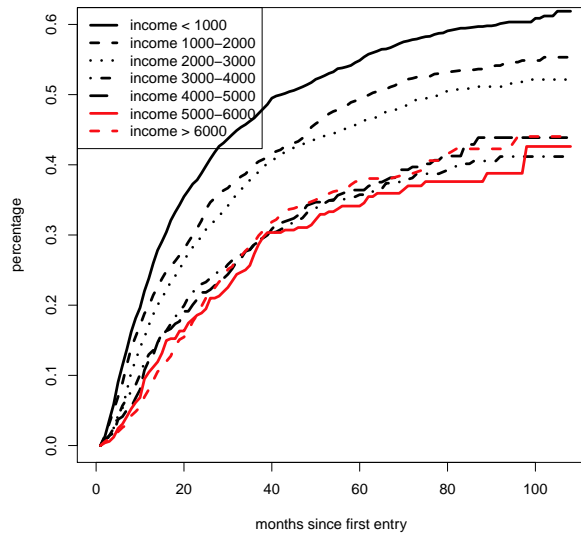
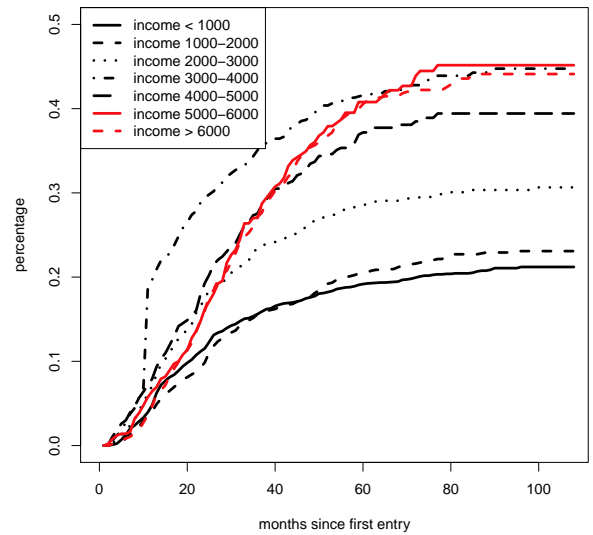Figure 1: Kaplan-Meier estimates of probability to stay in NL, by income group

Figure 2: Non-parametric survival rate and cumulative incidence functions

4   Simple duration analysis

We rely on duration analysis in our estimation of return migration for several reasons. First, duration analysis focuses on the timing of the return decision and not just on whether it occurred. A duration model takes into account such a change in intensity to leave. Second, along with the migration decisions, other relevant characteristics of the individuals may also change over time, such as the labor market status and migrant's income. Duration models allow us to include such time-varying covariates. Third, it is hardly ever possible to observe migration decisions over the whole life time of a migrant. The knowledge that the immigrant has been in the host country from his entry time up till the end, however, contains valuable information, and duration models allow for such right censoring as well as left truncation.

We assume that the conditional hazard follows a mixed proportional hazard model, given by products of baseline hazards (measuring duration dependence) and functions of observed time-varying characteristics x and unobserved characteristics v:

$$\theta(t|x(t)v) = v\lambda_0(t)\exp\big(x(t)\beta\big). \tag{1}$$

where $\lambda_0(t)$ represents the baseline intensity, that is, the duration dependence of the intensity common to all individuals.

If a migrant is administratively removed at duration t_a and the last observed change for this migrant occurred at duration $t_1 < t_a$, the contribution to the likelihood (of the out-migration) of this migrant is the probability of survival till t_1 times the probability that the migrant left the country between t_1 and t_a. The latter is equal to the survival from t_1 until t_a given survival.

Let a_i indicate whether the emigration of migrant i was due to an administrative removal ($a_i = 1$). For an administratively removed migrant we introduce two different event dates: t_i^a is the administrative removal date and $t_i^1 < t_i^a$ is the date of the last recorded change in any of the characteristics of migrant i before t_i^a.

We have data for $i = 1, \ldots, n$ immigrants entering the Netherlands in our observation window. We have the indicators Δ_i denoting that the migration spell is uncensored. Thus the likelihood contribution of migrant i conditional on the unobserved heterogeneity v is,

$$L = \prod_{i=1}^{n} \int \left\{ \left[\theta\big(t_i|x(t_i), v\big)^{\Delta_i} \exp\Big(-\int_0^{t_i} \theta\big(\tau|x(\tau), v\big)\, d\tau\Big) \right]^{(1-a_{ik})} \right.$$
$$\left. \cdot \left[\exp\Big(-\int_0^{t_i^1} \theta\big(\tau|x(\tau), v\big)\, d\tau\Big) - \exp\Big(-\int_0^{t_i^a} \theta\big(\tau|x(\tau), v\big)\, d\tau\Big) \right]^{a_i} \right\} dG(v) \tag{2}$$

where we assume that the unobserved heterogeneity follows a discrete distribution with three points of support, (v_1, v_2, v_3) and $\Pr(V = v_1) = p_1$, $\Pr(V = v_2) = p_2$, $\Pr(V = v_3) = 1 - p_1 - p_2$.[1]

4.1   Results of simple duration model

We assume a piecewise constant baseline intensity on seven intervals (at three and six months and at one, two, three and five years and beyond five years). The covariates included in the model refer to demographics (gender, age, marital status and age of children), country of origin's GDP per capita and GDP growth rate, and individual labor market characteristics (monthly income and industry sector). The individual's labor market history is also included.

We control for business cycle conditions by including the national unemployment rate, both at the moment of first entry to the country and the time-varying monthly rate. The unemployment rate at entry captures the 'cohort effect' of migrants, while the current varying unemployment rate captures the impact of the business cycle on the intensity to leave.

Table 2 presents the results for a proportional hazard model and a mixed proportional hazard model. We discuss the most relevant results. The income of migrants in the Netherlands has a U-shaped effect on the intensity to leave as both immigrants with low and high income leave faster. It is interesting to note that those with the lowest income (less than € 1000) have the highest probability of leaving.

Self-employed migrants have a stronger attachment to the Netherlands. Self-employment may imply a risky investment, which increases the ties to the country. It seems that those migrants are rather good in setting up a new business. House owners are, not surprisingly, less prone to leave. More migration experience makes the migrants more mobile internationally, see DaVanzo (1983).

Home country conditions seem to play an important role in return. Immigrants from poorer country of origin are less likely to leave, yet positive economic growth at home triggers return migration. High national unemployment rates, however, do lead to an increase in the departure of labor migrants. From the baseline duration dependence, we can conclude that the intensity to leave is low for the first three months in the country, then increases to a high for two years and then slowly decreases.

Those non-participating and having no income are more likely to leave but those unemployed on benefits are less prone to returning. These estimates should be interpreted with care as changes in the labor market status might be correlated with migration moves. If such selectivity exists it will bias the estimates of the effect of labor market changes on the migration intensity. Bijwaard

---

[1]We estimate $v_1 = \exp(a_1)$, $v_2 = \exp(a_2)$, $v_3 = \exp(a_3)$ and (q_1, q_2) with $p_1 = e^{q_1}/(1 + e^{q_1} + e^{q_2})$ and $p_2 = e^{q_2}/(1 + e^{q_1} + e^{q_2})$.

Table 2: Estimation results simple (M)PH model (with correction for administrative removal)

	PH		MPH	
Female	-0.263^{***}	(0.030)	-0.291^{***}	(0.033)
married	-0.089^{***}	(0.027)	-0.135^{***}	(0.031)
divorced	-0.409^{***}	(0.108)	-0.458^{***}	(0.116)
# of children	-0.268^{***}	(0.012)	-0.287^{***}	(0.011)
On benefit (unemployed)	-0.027	(0.104)	0.002	(0.102)
Non-participation	1.583^{***}	(0.045)	1.715^{***}	(0.048)
self-employed	-2.455^{***}	(0.287)	-2.488^{***}	(0.273)
income $<$ 1000	0.797^{***}	(0.066)	0.857^{***}	(0.066)
income 1000–2000	-0.215^{***}	(0.052)	-0.193^{***}	(0.052)
income 3000–4000	0.241^{***}	(0.055)	0.244^{***}	(0.056)
income 4000–5000	0.383^{***}	(0.068)	0.388^{***}	(0.069)
income 5000–6000	0.420^{***}	(0.082)	0.430^{***}	(0.084)
income $>$ 6000	0.600^{***}	(0.055)	0.611^{***}	(0.057)
repeated employment	-0.812^{***}	(0.046)	-0.945^{***}	(0.055)
Unemployed before	-0.204^{**}	(0.095)	-0.265^{**}	(0.107)
NP before	0.341^{***}	(0.056)	0.378^{***}	(0.062)
ln(GDPPC)	-0.086^{***}	(0.011)	-0.093^{***}	(0.012)
gdp growth	0.022^{***}	(0.003)	0.023^{***}	(0.003)
National UR	0.027^{**}	(0.014)	0.016	(0.015)
UR at entry	-0.327^{***}	(0.047)	-0.347^{***}	(0.050)
α_2 (3-6 mos)	1.215^{***}	(0.112)	1.230^{***}	(0.114)
α_3 (6-12 mos)	2.117^{***}	(0.099)	2.159^{***}	(0.106)
α_4 (12-24 mos)	2.455^{***}	(0.098)	2.552^{***}	(0.108)
α_5 (24-36 mos)	2.502^{***}	(0.100)	2.654^{***}	(0.112)
α_6 (36-60 mos)	2.485^{***}	(0.100)	2.721^{***}	(0.115)
α_7 ($>$ 60 mos)	2.399^{***}	(0.106)	2.727^{***}	(0.123)

Age, sector, entry year and country dummies are also included in the estimation.
$^{+}p < 0.10$; $^{**}p < 0.05$ and $^{***}p < 0.01$

et al. (2013) address this issue to obtain the causal effects of labor market changes on the return migration intensity by using a 'timing-of-events' method. In this paper the focus is on the impact of income, which depends on the labor market status, on the return migration intensity and not on the labor market changes itself. We therefore proceed with a method that takes this selectivity and the endogeneity of earned income into account.

5   A competing risks model

Our interest in this paper is to examine whether high- or low- income migrants return faster whilst controlling for the endogeneity of the labor market status and earned income, which impacts and the return migration process. We are interested, per se, in the labor market and the migration dynamics, the timing of the transitions and the time between transitions. Since we observe immigrants from the time they enter to the time they leave or till the end of our observation window, and since we focus on those employed immigrants at entry (after 3 months), an immigrant potentially faces different risks of exiting his/her first state of employment and multiple durations. Hence we use a competing risks model where there are several exit states. We define four states as follows:

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 (left the host country; i.e., returned)

These states are mutually exclusive and exhaust all possible destinations. A migrant may leave a state $j = 1, \ldots, 3$ (we ignore repeated immigration) for any of the other destination states, i.e. for $j = 1$ the destination states are $k = 2, 3, 4$, for $j = 2$ $k = 1, 3, 4$ etc. We view the migrant behavior as a semi-Markov process with individuals moving between the first three states and abroad as an absorbing state.

We use a competing risks model hazard model for each origin-destination pair. We define the random variables T_{jk} that describe the time since entry in 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 left 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[2], i.e. $\lambda_{0jk}(t) = \sum_{h=1}^{H} e^{\alpha_{jkh}} I_h(t)$ with $I_h(t) = I(t_{h-1} \leq t < t_h)$ 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)

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

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 are observed. The other durations are censored, in the sense that all is known that their realizations exceed $\tilde{T}_j$. If for individual i we observe M_{ijk} j to k transition spells, at sojourn times $t_1, \ldots, t_M$, then the likelihood contribution of these M_{ijk} transitions is:

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

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 depends on the labor market status, with by definition zero income in the non participation state, and the time spend t in this state

$$\ln W(t) = \xi_0 + \sum_{h=1}^{H} \xi_h I_h(t) + \xi_2 x(t) + \epsilon(t) \quad (6)$$

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

$$L^w\big(W(1), \ldots, W(t)|x(1), \ldots, x(t), v_w\big) = \prod_{s \leq t} \phi\left(\frac{\ln W(s) - \xi_0 - \sum_{h=1}^{H} \xi_h I_h(s) - \xi_2 x(s) - v_w}{\sigma_\eta}\right) \quad (7)$$

---

[2]It is not necessary that each baseline intensity changes at the same durations. 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 $\alpha_{jkh} = \alpha_{jkh+1}$.

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

For the sake of parsimoniousness, 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 discrete support $(V_1, \ldots, V_M)$ and $p_m = \Pr(V = V_m)^3$. It is important to note that the V_m's are vectors with $V_m = (V_{12m}, V_{13m}, V_{14m}, V_{21m}, V_{23m}, V_{24m}, V_{31m}, V_{32m}, V_{34m}, V_{wm})'$ including all the possible transitions and the random components of the income equations.

The complete likelihood function for each individual is

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

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

5.1 Results of the competing risks model

The number of vectors of support is chosen to be $M = 3$. Table 3 presents the estimated income coefficients of all the transitions involved. [4] However, the interpretation of the coefficients in a competing risks model requires caution.[5] A particular covariate, say x_l, can appear in several intensities. In such a case the vectors β_{ljk} convey little information about the effect of the covariate on the probability to exit from origin j to destination k. The reason is that the exit probability depends not only on the intensity of making a transition to k but also on the transition intensities to all other states.

For this reason we only mention the main finding of the income effect on the transition intensities.[6] When a migrant is employed income has a U-shaped effect on return migration (transition to abroad), reflecting what we have found for the simple duration model. The transition to unemployment is negatively related to the income while employed and, the income effect of the transition to non-participation is U-shaped. However, migrants can leave the country after some period of unemployment/non-participation, or after more intermediate states. The multi-state competing risk framework takes this into account, but makes the interpretation of the coefficients difficult.

---

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

[5]Note that in a standard mixed proportional hazard (MPH) model, the interpretation of the coefficients is also not straightforward. In an MPH model, the regression coefficient of covariate x_l is only defined conditionally on the unobserved heterogeneity.

[6]Appendix A reports the estimated correlation structure of unobserved heterogeneity across transition probabilities and shows the difference in predictions of the simple duration model and the competing risks model.

Table 3: Income coefficient estimates for correlated competing risks model

| | from EMPLOYED | | |
	Unemployed	Non-participation	Abroad
income 0- 1000	1.027***	0.750***	0.834***
	(0.100)	(0.041)	(0.074)
income 1000-2000	0.132	0.239***	−0.219***
	(0.083)	(0.032)	(0.054)
income 2000-3000	−	−	−
income 3000-4000	−0.164	−0.009	0.268***
	(0.141)	(0.045)	(0.057)
income 4000-5000	−0.406[+]	0.091	0.450***
	(0.226)	(0.057)	(0.069)
income 5000-6000	−0.175	0.118[+]	0.525***
	(0.268)	(0.070)	(0.084)
income > 6000	−0.456**	0.170***	0.762***
	(0.206)	(0.071)	(0.060)
	from UNEMPLOYED		
	Employed	Non-participation	Abroad
income < 1000	−0.139**	−0.367***	−0.801***
	(0.068)	(0.084)	(0.225)
income >1000	−	−	−

Income coefficients from non-participation are absent because all migrants in non-participation has zero income. $^{+}p < 0.10$; $^{**}p < 0.05$ and $^{***}p < 0.01$

5.2 Transition probability in multi-state models

The difficulty in interpreting the covariate effects also arises in many other non-linear models, such as the multinomial logit and probit models (see e.g. Cameron and Trivedi (2005), chapter 15). The results of such models are, therefore, usually reported in terms of the marginal effects on the probability of interest. Thomas (1996) and Kyyrä (2009) argue that a similar practice is useful in the context of competing risks models. Although the marginal effects eliminate much of the confusion in the interpretation of the results from competing risks models, they have rarely been computed. A drawback is that in general the marginal effects have no analytical solution, making their computation demanding and statistical inference difficult. Kyyrä (2009) shows that simple closed form solutions exist for the competing risks models with piecewise constant baseline hazards and discrete unobserved heterogeneity, exactly the model formulation we assume.

To look further ahead, we need to take all the transitions into account. In a multi-state model, migrants can return to the state they were once before. An employed migrant may, as we observe in our data, first become a non-participant before he leaves the country. Another possible route to leave the country is through unemployment and non-participation. It is even possible that the migrant, after a period of unemployment, returns to work and then leaves the country. The transition probability, which is the probability to be in a particular state given the time since entry, takes all the possible intermediate transitions into account. Dabrowska et al. (1994) describe how we can derive these transition probabilities for the semi-Markov model we use.

The transition probability from state j to state k after a duration t (where t is now the time since the migrant entered the host) is formed by adding all possible intermediate transitions that start in j and end in k at time t. First consider the migrants who do not make a transition in $(0, t)$, thus $j = k$. Those individuals remain in j till t, they are the migrants who remain working. The probability that the employed remain working is equal to the total survival of the employed, $S_j(t)$, i.e.

$$S_j(t|\overline{X}_{jk}(t)) = \Pr(\tilde{T}_j \geq t) = \prod_{l \neq j} \int \exp\left(-\Lambda_{jl}(t|\overline{X}_{jl}(t), V_{jk})\right) dG_{jl}(V_{jl}) \tag{9}$$

Next we have the migrants who make one transition within a period t since they entered the country, say from employment to non-participation, and then remain in this state till the end of the period. The probability that a transition from j to k before t occurs and the migrants then remain in k is equal to

$$\int_0^t f_{jk}(u|\cdot) \cdot S_k(t - u) \, du$$

with $f_{jk}(t) = \partial F_{jk}(t)/\partial t$, the cumulative incidence function.[7] Conditional on unobserved heterogeneity the cumulative incidence can be expressed as

$$
\begin{aligned}
F_{jk}(t|\overline{X}_{jk}(t), V_{jk}) &= \Pr\big(\tilde{T}_j \leq t, \text{destination} : k\big) = \int_0^t \lambda_{jk}(s|\overline{X}_{jk}(s), V_{jk}) S_j(s|\overline{X}_{jk}(s), V_{jk})\, ds \\
&= \sum_{h=1}^{H} \pi_{jk}^h(X|V_{jk})\Big[S\big(t_{h-1}|\overline{X}_{jl}(t), V_{jk}\big) - S\big(t_h|\overline{X}_{jl}(t), V_{jk}\big)\Big] J_h(t) \qquad (10) \\
&\quad + \sum_{h=1}^{H} \pi_{jk}^h(X|V_{jk})\Big[S\big(t_{h-1}|\overline{X}_{jl}(t), V_{jk}\big) - S\big(t|\overline{X}_{jl}(t), V_{jk}\big)\Big] I_h(t)
\end{aligned}
$$

where $\pi_{jk}^h(X|V_{jk})$ denotes the probability of exit from j to k in interval $[t_{h-1}, t_h)$ conditional on exiting and $S(t_{h-1}|\cdot) - S(t_h|\cdot)$ is the probability of exiting j during the interval $[t_{h-1}, t_h)$. Integrating the correlated (over $9 \cdot M$) discrete unobserved heterogeneity we obtain

$$
F_{jk}(t|\overline{X}_{jk}(t)) = \sum_q \Pr(V_j = V_j^q) F_{jk}(t|\overline{X}_{jk}(t), V_j^q) \qquad (11)
$$

with $V_j = \{V_{jk}, k \neq j\}$ and the sum is over all possible realizations of V_j (27 in our application with a 3-point discrete unobserved heterogeneity distribution and three exit states).

Some migrants may, after first making a transition from employment to non-participation, end up abroad. The probability of making a transition from j to k within a period t with one intermediate initial transition is

$$
F_{jk}^{(2)}(t|\cdot) = \int_0^t \sum_{m=1}^{4} F_{jm}(u|\cdot) \cdot f_{mk}(t-u|\cdot)\, du
$$

with the cumulative incidence from j to j, $F_{jj}(t|\cdot) = 0$. Then, the probability that a migrant who made these two transitions and who remains in state k till t is

$$
\int_0^t f_{jk}^{(2)}(u|\cdot) S_k(t-u)\, du,
$$

with $f_{jk}^{(2)}(u|\cdot) = \partial F_{jk}^{(2)}(t)/\partial t$. This reasoning is repeated for any number of intermediate transitions from state j to state k Thus, the transition probability, i.e. the probability to be in k starting in j after a duration t is

$$
P_{jk}(t|\cdot) = S_j(t|\cdot) \cdot I(j=k) + \sum_{p \geq 1} \int_0^t f_{jk}^{(p)}(u|\cdot) S_k(t-u)\, du \qquad (12)
$$

where $f_{jk}^{(p)}(t) = \partial F_{jk}^{(p)}(t)/\partial t$ and

$$
F_{jk}^{(p)}(t|\cdot) = \int_0^t \sum_{m=1}^{4} F_{jm}^{(p-1)}(u|\cdot) \cdot f_{mk}(t-u|\cdot)\, du
$$

---

[7]The cumulative incidence function is also known under the name 'subdistribution function'. This name reflects that the cumulative probability to make the $j-k$ transition remains below one, $F_{jk}(\infty|\cdot) < 1$. Note that $\sum_{k \neq j} F_{jk}(t|\cdot) = 1 - S_j(t|\cdot)$.

In this paper, we use data on labor migrants only and are interested in return migration. By definition, all labor immigrants to the Netherlands are employed at entry. Thus, we are only interested in the transition probability from employment to abroad, the return migration probability. After estimating the competing risks model for all the possible transitions, we will derive the path of the return migration probability for the reference individual and discuss the impact of income differences on this probability.

5.3 Comparing results with simple duration model

Note that for a simple (one state) duration model, the return migration probability is the cumulative density function, the probability to experience the event after a duration t. We calculate for both the simple and the correlated competing risk model (ccrm) the return migration probability for the recent labor migrants (from employed). Figure 3 presents these return migration probabilities for the reference migrant, a single male aged 30 to 35, employed in the trade sector from a country with a GDP per capita of $2000 who entered the Netherlands in 2001 and lives in a rental house. Note that the simple model underestimates the return migration of the migrants. Five years after their arrival 62% (33% for the simple model) of the labor migrants have left the country. After ten years the percentage of migrants that have left the country has increased to 83% (55% according to the simple model).

Figure 4 presents the marginal (as a function of the time since entry) effect of initial income on the return migration probability both for the CCRM model and for the simple model. First we observe that the simple model overestimates the long run income effects on the return migration probability. When taking labor market changes into account, low-income migrants have a 9% higher probability to leave (this difference remains rather constant after five years since entry). Low-income migrants have a much higher probability of becoming unemployed or non-participating and migrants are more prone to leave when not employed. The simple model does not take this relation between the labor market status and migrant income into account.
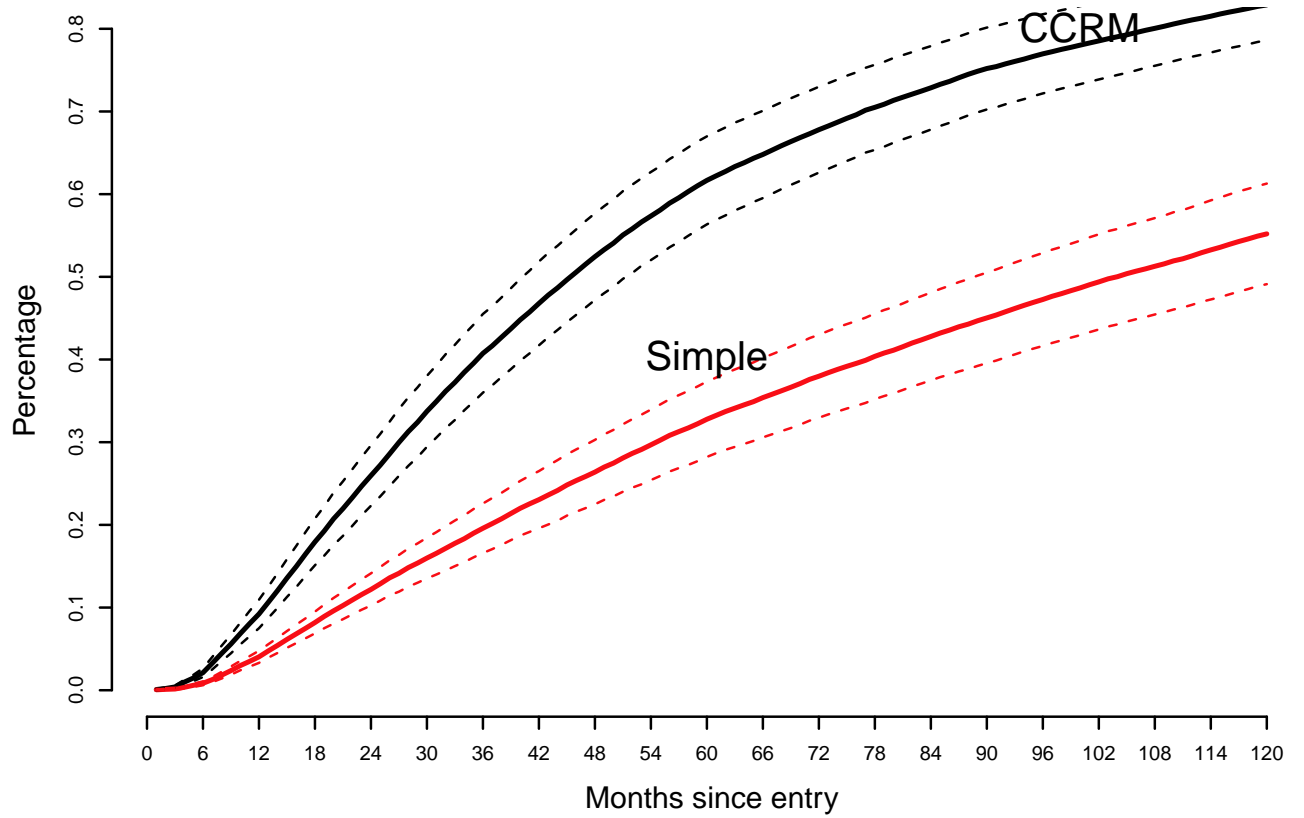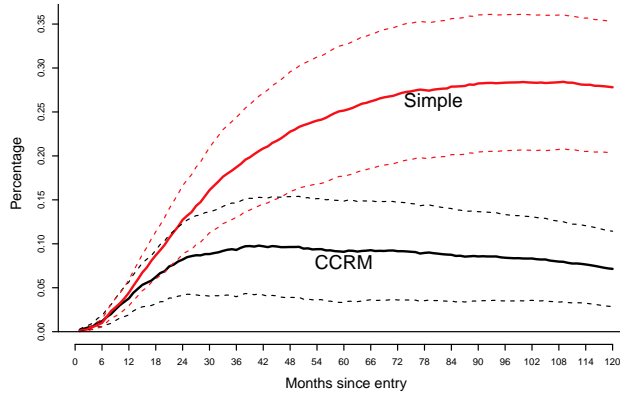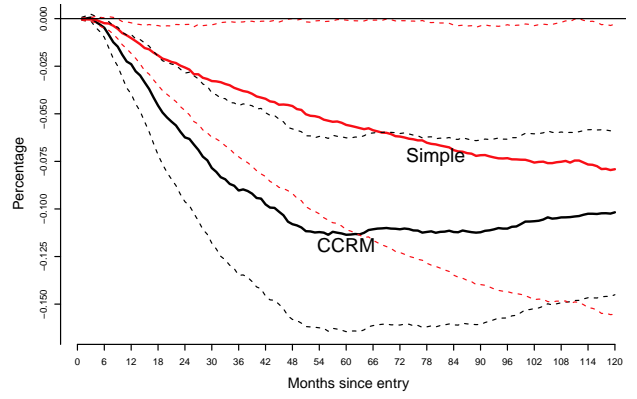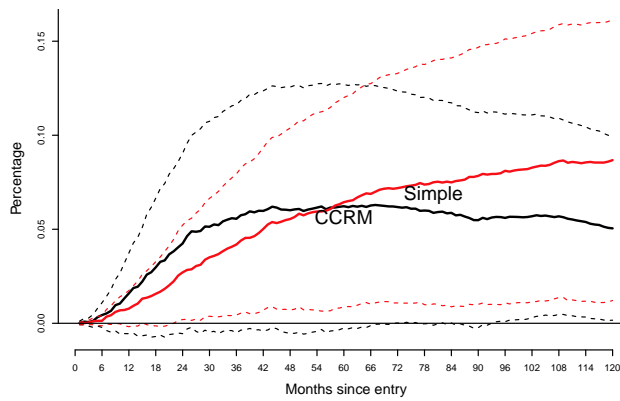
Figure 3: Probability of return with 95% confidence bands (reference individual)
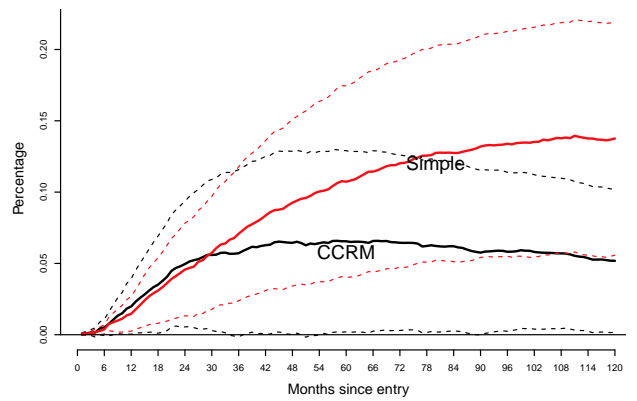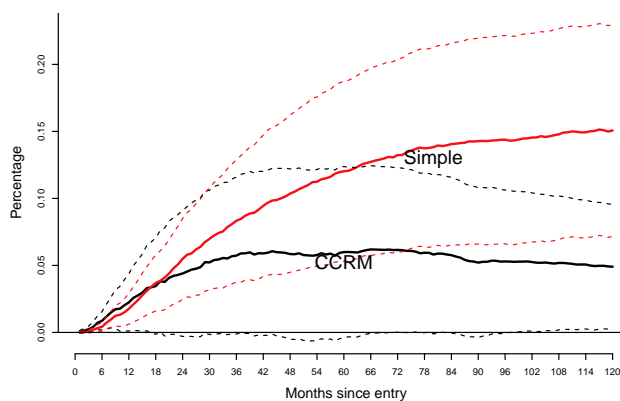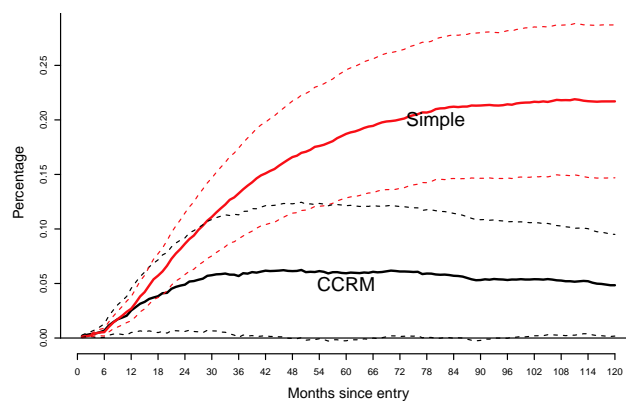
Figure 4: Marginal effect of initial income on return

Table 4: Marginal income effect on return probability by duration in NL

	Inc 1	Inc 2	Inc 3	Inc 4	Inc 5	Inc 6	Inc 7
			Initial income: from low to high				
1 year	0.0378***	−0.0239**	-	0.0159	0.0198$^+$	0.0224$^+$	0.0253**
2 year	0.0820***	−0.0624***	-	0.0424	0.0496$^+$	0.0440	0.0489$^+$
3 year	0.0930***	−0.0903***	-	0.0556	0.0571	0.0573	0.0568$^+$
4 year	0.0964***	−0.1080***	-	0.0601	0.0644$^+$	0.0585	0.0614
5 year	0.0909***	−0.1136***	-	0.0622	0.0655$^+$	0.0600	0.0594
6 year	0.0916***	−0.1107***	-	0.0618	0.0645$^+$	0.0616$^+$	0.0610$^+$
7 year	0.0875***	−0.1121***	-	0.0587$^+$	0.0620$^+$	0.0577	0.0571
8 year	0.0841***	−0.1103***	-	0.0560$^+$	0.0582$^+$	0.0530	0.0532
9 year	0.0797***	−0.1044***	-	0.0569$^+$	0.0570$^+$	0.0518$^+$	0.0527$^+$
10 year	0.0714***	−0.1018***	-	0.0505$^+$	0.0518$^+$	0.0490$^+$	0.0484$^+$

Inc 1: Monthly income $< €1000$; Inc 2: Monthly income $€1000$-$€2000$; Inc 3: Monthly income $€2000$-$€3000$; Inc 4: Monthly income $€3000$-$€4000$; Inc 5: Monthly income $€4000$-$€5000$; Inc 6: Monthly income $€5000$-$€6000$; Inc 7: Monthly income $> €6000$. $^+ p < 0.10$; $^{**} p < 0.05$ and $^{***} p < 0.01$

Indeed, Table 4 summarizes the marginal effects of initial income on return probability by duration in the Netherlands. The U-shaped relationship between initial income and return is clear and also the lowest earners have the highest probability of return among all income groups regardless of their migration duration. There is evidence of failure leading to return migration as those with the lowest income have the highest probability for the first year, and about 30% more likelihood, to return compared with the next likely group (the top earners). In addition, the probabilities of return peak at about 3-4 years for the lowest income group, whilst for the other income groups peaks a bit later at about 4-6 years. The gap in the intensity of return between the lowest- and highest-income group does not decline over time. Although we still find a U relationship between initial income and return intensity, there is no significant difference among the high income groups, earning above the average income, income groups 4-7.

6 Microsimulation

The return migration probability gives the probability that a labor migrant is abroad after a given time since the migrant entered the country. It takes the full dynamics into account. However, this transition probability hides the information on how an individual reached a certain state. Many relevant indicators of the paths of the immigrants on the host labor market, e.g. the average length of an unemployment spell, cannot be derived analytically. In this section we provide these indicators on the basis of microsimulations. These simulations use the estimated parameters of the correlated competing risks model and the observed entry into the Netherlands as input.

This simulation is based on a synthetic cohort of labor migrants, all entering at the same time. The synthetic cohort consists of 50,000 migrants, for which the distribution of the start population of migrants equals the observed entry distribution. For each simulation round, we draw a vector of parameter estimates assuming that the estimated coefficients are normally distributed around the point estimates with a variance-covariance matrix equal to the estimated one. Then, on a monthly basis, we simulate the transitions for each member of the synthetic cohort using the implied transition intensities. If the simulated migrant becomes unemployed, we use the transition intensity from unemployment, and similarly for a non-participating migrant and a migrant abroad. In the simulations the exogenous explanatory factors remain at their initial value. The (endogenous) value of the income of the migrant increases over the length of the time spent in employment using the implied income increase obtained from the estimated ccrm. We use the evolution of the labor-migration path, the history of all occurrences of labor market and migration states, of each individual member in the (dynamic) simulation. Thus, if a (simulated) migrant finds a job again after some period of unemployment, we take the effect of the labor market experience into account. We simulate the labor-migration path for ten years, and in the end we save the whole simulated migrant history. We repeat the simulations 100 times.

Table 5: Simulation results for 10 years

	Initial income: from low to high						
	Inc 1	Inc 2	Inc 3	Inc 4	Inc 5	Inc 6	Inc 7
Average time in NL	47.2	67.4	56.7	50.7	50.4	50.8	50.6
Fraction of time in NL employed	65.8%	79.6%	78.4%	77.7%	78.0%	78.0%	77.9%
Fraction of time in NL unemployed	6.3%	3.4%	3.3%	2.3%	2.1%	2.3%	2.2%
Fraction of time in NL no income	27.9%	17.0%	18.3%	20.0%	19.9%	19.8%	19.9%
Fraction unemployed within 10 years	15.7%	12.3%	10.0%	8.6%	8.2%	8.5%	8.3%
Fraction no-income within 10 years	69.5%	65.7%	58.9%	56.6%	56.0%	56.5%	56.3%
Average # employment spells	1.52	1.38	1.32	1.28	1.27	1.28	1.28
Average spell length if employed	20.4	38.7	33.8	30.8	30.9	30.9	30.9
Average # unemployment spells	0.42	0.25	0.18	0.15	0.15	0.18	0.15
Average spell length if unemployed	7.2	9.3	10.2	7.6	7.4	7.7	7.5
Average # no income spells	0.84	0.73	0.65	0.62	0.62	0.62	0.62
Average spell length if no income	15.7	17.0	16.1	16.3	16.2	16.1	16.2

Inc 1: Monthly income < € 1000; Inc 2: Monthly income € 1000-€ 2000; Inc 3: Monthly income € 2000-€ 3000; Inc 4: Monthly income € 3000-€ 4000; Inc 5: Monthly income € 4000-€ 5000; Inc 6: Monthly income € 5000-€ 6000; Inc 7: Monthly income > € 6000;

Table 5 presents some labor market and migration indicators and Table 6 presents the average paths of the migrants on the labor market. Both these simulations results are differentiated by initial income level. It is obvious that the low income migrants spend more time unemployed and non-participating and less time employed. The distinction between the high income groups levels off, though the difference between the lowest, middle and high groups is still apparent. Almost 16% of the lowest-income migrants have been unemployed within ten years of arrival. When they become unemployed they are unemployed for slightly more than half a year. However, the migrants in the lowest-income group also stay less than a little over one year in the country. Still, more than 8% of the high-income groups (groups 4-7) have been unemployed within ten years in the Netherlands and stay on average unemployed for 7.5 months. More than half of the migrants experience a period with no income (70% for the lowest-income group). On average they are without income for about one year and four months. From Table 6 we can derive that the majority of these migrants without income remain in the country after their job has finished, as 37% to 51% of the labor migrants returning after first experiencing a period of no income. Another interesting fact from Table 6 is that only a small portion of the migrants remains employed for the full (simulated) ten-year period. For the lowest-income group only 0.3% of the migrants remains employed for the whole ten years. The lower-income groups leave the country more often after one (or more) labor market changes. Interestingly, the most common path for the lowest-income group is non-participation before return (45%), whilst for the highest-income groups is leaving straight from employment (45%).

Table 6: Labour market paths

	Initial income: from low to high						
	Inc 1	Inc 2	Inc 3	Inc 4	Inc 5	Inc 6	Inc 7
Most common paths							
% Employed for 10 years	0.3%	7.3%	3.6%	2.3%	2.3%	2.4%	2.5%
% Employed-abroad	26.9%	24.2%	35.5%	39.8%	40.4%	39.7%	39.9%
Average employment duration	23.6	35.6	32.4	30.2	29.8	29.9	29.7
% Employed-NP-abroad	40.6%	36.8%	33.6%	33.0%	32.8%	32.8%	32.8%
Average employment duration	17.5	28.6	25.8	23.7	23.9	23.8	23.6
Average no income duration	20.4	19.4	19.7	20.0	20.1	20.0	20.2
% Employed-NP-employed	2.5%	8.4%	5.5%	3.8%	4.0%	4.1%	4.1%
Average 1st employment duration	28.4	39.2	36.8	34.5	36.3	35.1	37.1
Average no income duration	14.0	10.1	11.3	11.9	11.9	12.1	11.1
Average 2nd employment duration	78.6	71.7	72.9	74.6	72.8	73.8	72.8
% Emp-NP-emp-abroad	10.0%	5.9%	7.9%	9.1%	9.0%	9.1%	9.0%
Average 1st employment duration	16.5	24.1	22.4	20.9	20.9	20.9	21.4
Average 1st no income duration	8.0	7.6	7.8	7.7	7.8	7.8	7.8
Average 2nd employment duration	31.9	37.2	35.0	34.7	33.8	34.9	33.4
Most common paths ending abroad[a]							
% Employed-abroad	29.8%	33.2%	42.7%	45.1%	45.7%	45.1%	45.4%
% Employed-NP-abroad	45.0%	50.5%	40.5%	37.5%	37.2%	37.3%	37.3%
% Emp-NP-emp-abroad	11.0%	8.0%	9.5%	10.4%	10.2%	10.3%	10.3%
% Emp-unemp-emp-abroad	1.9%	1.1%	1.3%	0.8%	0.8%	0.9%	0.8%

Inc 1: Monthly income < €1000; Inc 2: Monthly income €1000-€2000; Inc 3: Monthly income €2000-€3000; Inc 4: Monthly income €3000-€4000; Inc 5: Monthly income €4000-€5000; Inc 6: Monthly income €5000-€6000; Inc 7: Monthly income > €6000.
[a] Percentage of all paths ending abroad.

7 Specific Countries

7.1 Descriptive statistics

Given the potential variation between countries of origin, we focus here on five main countries of labor immigration to the Netherlands, namely India, Turkey, China, South Africa and Morocco. As seen in Table 7, almost 19% of recent labor immigrants came from India and 10% from China. labor immigrants from Turkey represented 11% and those from Morocco were only 3% as the majority of immigrants from these two countries tend be family migrants rather than labor immigrants. Finally, 8% of immigrants came from South Africa. The distribution of income group shows that labor migrants from Morocco and China more often start with low paying jobs, while Indian and South-African migrants are overrepresented in high-paying jobs. South African migrants are more often female, due to a Dutch policy to attract nurses from that country, while only a few Indian migrants are female. Indian labor migrants are also younger and more often single. Given the small proportion of our immigrants earning above € 3000, we aggregate the high income groups together in the analysis below.

Table 7: Descriptive statistics at entry: LDC labor migrants

	Country of origin				
	India	Turkey	China	South-Africa	Morocco
Female	11.4%	17.3%	19.6%	28.9%	20.7%
Single	73.6%	68.9%	63.8%	64.3%	64.0%
Married	26.3%	29.9%	35.5%	34.5%	33.1%
av. age	29.0	31.9	30.5	31.6	32.4
GDP pc.	$787	$5228	$1866	$3722	$1524
GDP growth	8.4%	4.2%	11.6%	3.9%	4.0%
	Income distribution				
< € 1000	20.6%	33.1%	42.5%	21.7%	56.5%
€ 1000- € 2000	17.5%	27.2%	35.2%	20.6%	25.4%
€ 2000- € 3000	25.1%	24.5%	10.7%	21.5%	9.6%
€ 3000- € 4000	17.5%	7.1%	4.3%	13.1%	3.5%
€ 4000- € 5000	8.9%	2.9%	2.1%	6.6%	1.2%
€ 5000- € 6000	3.0%	1.2%	1.5%	3.9%	0.2%
> € 6000	7.5%	4.1%	3.8%	12.6%	3.7%
N=	3261	1851	1726	1309	492
%	18.7%	10.6%	9.9%	7.5%	2.8%

7.2 Transition probability in multistate models

Figure 5, which is based on the estimated competing risks model, shows the probability of return by duration of migration. Indians and Chinese show the highest probability of return whilst the
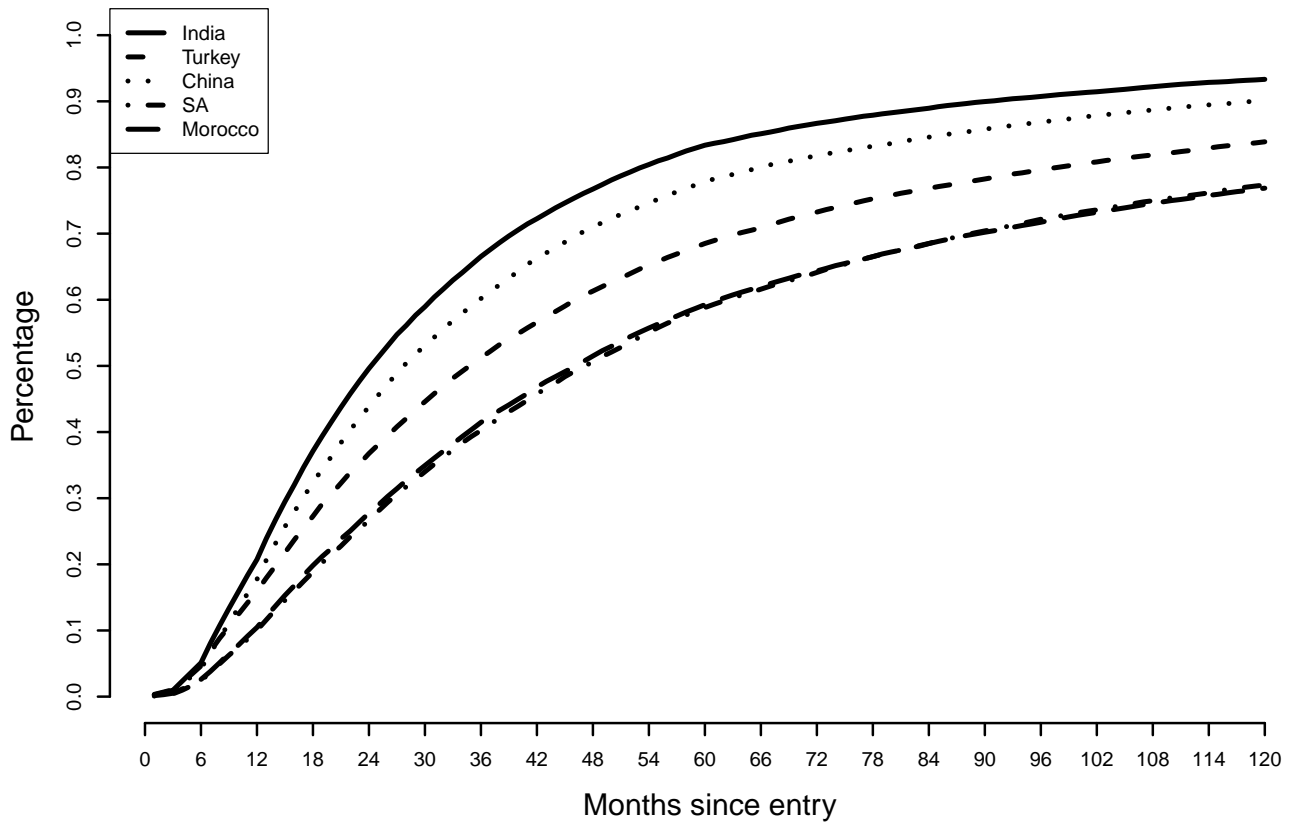
Figure 5: Probability of return by main country of origin (reference individual)

Moroccans and South Africans exhibit the lowest return. Figure 6 reveals that immigrants from India and China have, after controlling for labor market and earned income changes, the strongest income effect on return after two years, but after six years in the Netherlands, South Africans and Moroccans have the highest income effects.

In accordance with our results from Section 5.3, for each of our countries of origin we find consistently a U-shaped relationship between initial income and return, Figure 7, and that the lowest income group has the highest probability of return followed by the top earners whilst middle income groups have lower return probabilities. Although a U pattern is found for immigrants from each country, there is a variation in the impact of income on return among countries.
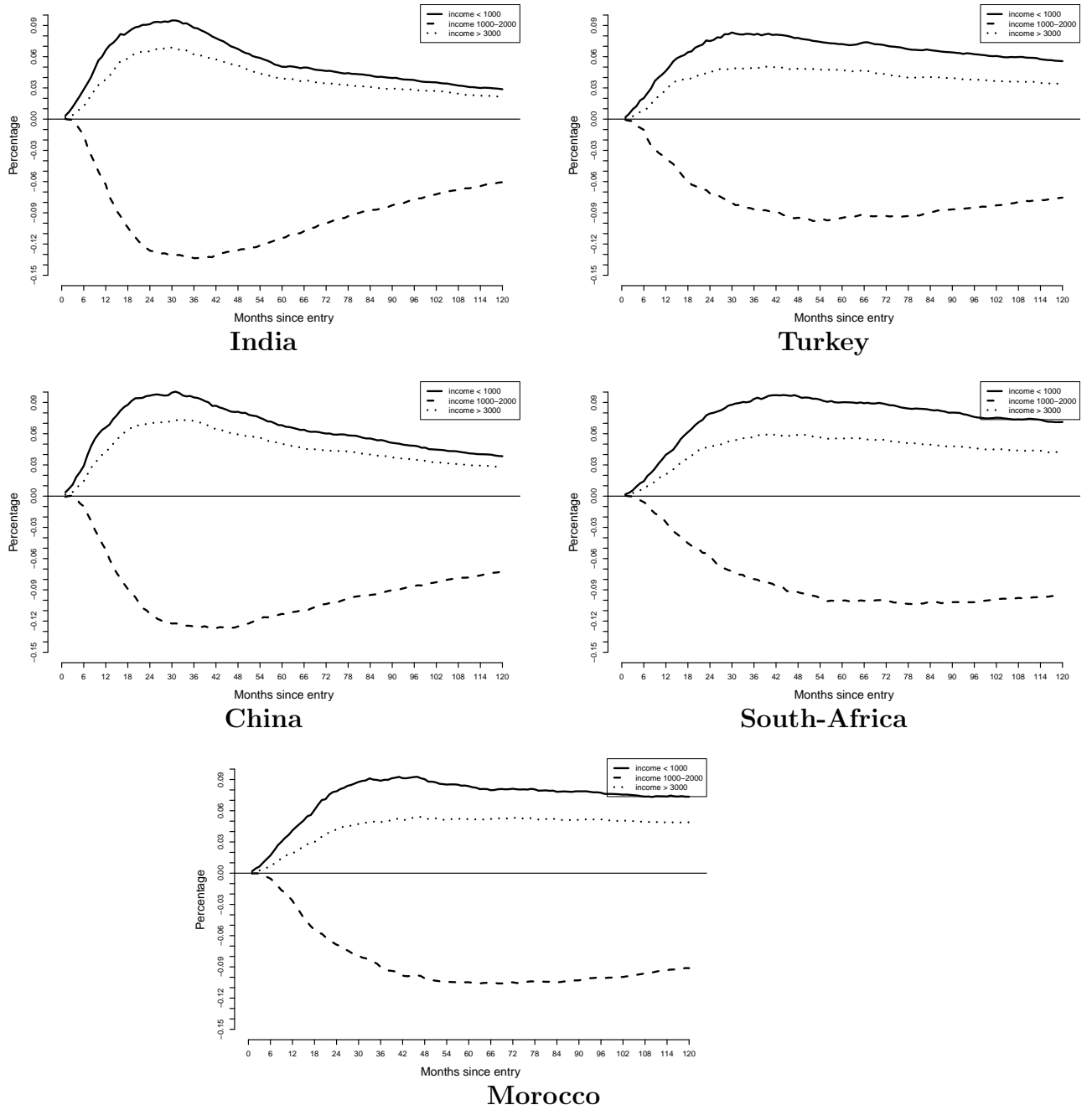
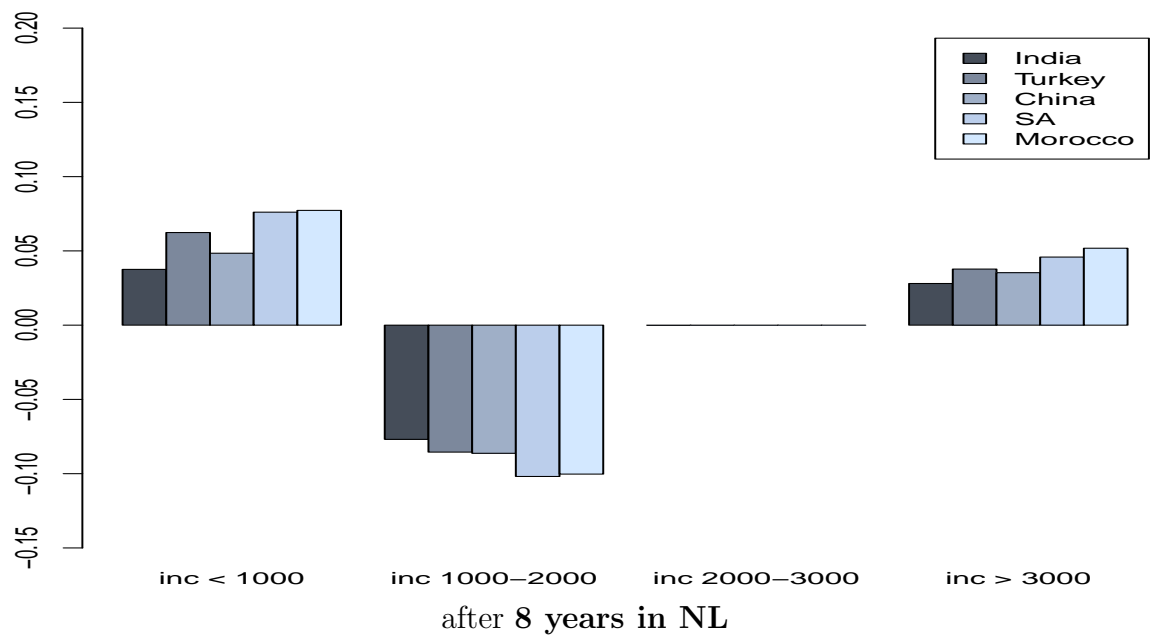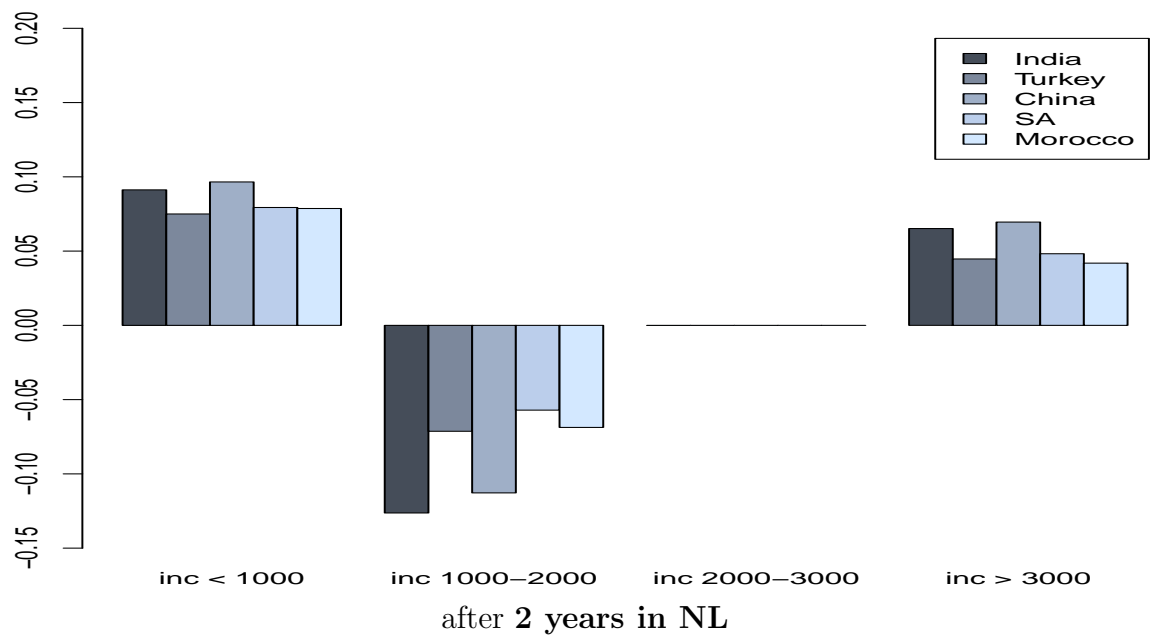Figure 6: Marginal effect of initial income on return, by country

Figure 7: Marginal effect of initial income on the probability abroad

A similar simulation as in Section 6 was carried out for each of the five countries separately. These simulation results also indicate the variations in the labor market path among immigrants from different countries controlling for their income group. For example, among the lowest-income group, 7% of Indians experience unemployment within 10 years in the Netherlands compared with 21% among Moroccans. On the other hand, looking at all Indian (Moroccans) immigrants, Table 8 shows that those with initially less than €1000 stay on average 29 (50) months in the Netherlands, whilst the high earners stay 31 (54) months.

Table 8: Simulation results for 10 years, main countries of origin

	Country of origin				
	India	Turkey	China	South-Africa	Morocco
Initial monthly income < €1000					
Average time in NL	29.3	41.9	33.5	50.5	50.4
Fraction of time in NL employed	78.5%	66.5%	77.5%	70.2%	65.4%
Fraction of time in NL unemployed	5.3%	8.3%	6.2%	8.7%	14.0%
Fraction of time in NL no income	16.1%	25.2%	16.3%	21.1%	20.6%
Fraction unemployed within 10 years	7.1%	12.8%	8.5%	15.9%	23.0%
Fraction no-income within 10 years	42.1%	65.5%	41.9%	61.5%	67.5%
Initial monthly income at entry €1000-€2000					
Average time in NL	47.0	59.0	52.4	69.4	69.5
Fraction of time in NL employed	88.8%	80.5%	87.9%	83.0%	80.7%
Fraction of time in NL unemployed	2.3%	4.2%	2.8%	4.3%	6.8%
Fraction of time in NL no income	8.9%	15.3%	9.3%	12.7%	12.4%
Fraction unemployed within 10 years	5.3%	10.0%	6.5%	11.6%	16.7%
Fraction no-income within 10 years	39.4%	60.4%	39.0%	55.6%	62.3%
Initial monthly income €2000-€3000					
Average time in NL	35.8	49.7	41.2	59.6	59.3
Fraction of time in NL employed	87.7%	79.5%	87.1%	82.4%	79.9%
Fraction of time in NL unemployed	2.3%	3.9%	2.8%	3.7%	6.6%
Fraction of time in NL no income	10.0%	16.5%	10.1%	13.9%	13.5%
Fraction unemployed within 10 years	3.7%	7.4%	4.9%	8.9%	13.2%
Fraction no-income within 10 years	33.5%	55.1%	33.2%	51.1%	57.7%
Initial monthly income at entry > €3000					
Average time in NL	31.1	45.0	35.6	54.1	53.9
Fraction of time in NL employed	87.3%	79.2%	87.2%	82.5%	80.9%
Fraction of time in NL unemployed	1.6%	2.9%	2.0%	2.8%	4.7%
Fraction of time in NL no income	11.1%	17.9%	10.8%	14.7%	14.4%
Fraction unemployed within 10 years	3.0%	6.5%	4.0%	7.4%	11.1%
Fraction no-income within 10 years	31.2%	52.9%	30.9%	49.1%	55.5%

8   Conclusion

The impact of income earned overseas is theoretically ambiguous with regard to return migration. Migrants would, on the one hand, like to extend their stay overseas as a response to higher wages; on the other hand, the gain from staying longer abroad decreases. As a consequence, higher wages abroad may have a positive or a negative effect on migration duration. In this paper, we estimate the impact of income earned in the host country on return migration of labor migrants. We use a four state correlated competing risks model to account for the strong dependence of the migrant's labor market status and earned income. In addition, we control for the changes in country of origin's economic growth and GDP per capita. For the analysis we use unique administrative panel data of recent labor immigrants from LDCs to the Netherlands.

The empirical results reveal that return intensities are U-shaped with respect to initial income with high intensity for low- and high- income groups and the lowest-income group exhibiting the highest return. We also find that ignoring the interdependence of labor market status and incomes earned leads to overestimating the long run impact of income differences on return. The fact that low income migrants return faster can be interpreted as a result of failure. On the other hand high earners leaving is due to them successfully meeting their target savings or acquiring their planned skills. Our results show that although the intensity of return varies by duration and country of origin, the U-shaped relationship between initial income and return is consistently found.

Our findings have important policy immigration implications. It is interesting to underscore that less successful immigrants return and thus the overconcern by host countries being burdened by welfare seekers is unfounded. Furthermore, the return of the more successful immigrant indicates that the concern by LDCs about the brain drain is exaggerated as migration might lead to brain circulation.

References

Bijwaard, G. E. (2009). Labour market status and migration dynamics. Discussion Paper No. 4530, 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 (2013). The impact of labour market dynamics on the return–migration of immigrants. *Review of Economics and Statistics forthcoming.*

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.

Borjas, G. J. and B. Bratsberg (1996). Who leaves? The outmigration of the foreignborn. *The Review of Economics and Statistics 78*, 165–176.

Cameron, A. C. and P. K. Trivedi (2005). *Microeconometrics: Methods and Applications.* Cambridge University Press.

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

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

Dabrowska, D. M., G. Wen, and M. M. Horowitz (1994). Cox regression in a Markov Renewal model: An applocation to the analysis of bone marrow transplant data. *Journal of the American Statistical Association 89*, 867–877.

DaVanzo, J. (1983). Repeat migration in the United States: Who moves back and who moves on. *The Review of Economics and Statistics 65*, 552–559.

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

Dustmann, C. (1997). Return migration, uncertainty and precautionary savings. *Journal of Development Economics 52*, 295–316.

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

Dustmann, C. and Y. Weiss (2007). Return migration: Theory and empirical evidence. *British Journal of Industrial Relations 45*, 236–256.

Galor, O. and O. Stark (1991). The probability of return migration, migrants' work effort, and migrants performance. *Journal of Development Economics 35*, 399–405.

Gibson, J. and D. McKenzie (2011). The microeconomic determinants of emigration and return migration of the best and brightest: Evidence from the Pacific. *Journal of Development Economics 95*, 18–29.

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.

Kyyrä, T. (2009). Marginal effects for competing risks models with piecewise constant hazards. *Oxford Bulletin of Economics and Statistics 71*, 539–565.

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.

Nekby, L. (2006). The emigration of immigrants, return, vs onward migration: Evidence from Sweden. *Journal of Population Economics 19*, 197–226.

Thomas, J. M. (1996). On the interpretation of covariate estimates in independent competing risks models. *Bulletin of Economic Research 48*, 27–39.

Yang, D. (2006). Why do migrants return to poor countries? Evidence from Philippine migrants' responses to exchange rate shocks. *The Review of Economics and Statistics 88*, 715–735.

A Correlation between the unobserved heterogeneity terms

Table A.1: Correlation between the unobserved heterogeneity terms

	v_{eu}	v_{en}	v_{ea}	v_{ue}	v_{un}
v_{eu}	–	0.899***	0.833***	−0.724***	−0.998***
v_{en}	0.899***	–	0.506	−0.953***	−0.870***
v_{ea}	0.833***	0.506	–	−0.221	−0.865***
v_{ue}	−0.724***	−0.953***	−0.221	–	0.680***
v_{un}	−0.998***	−0.870***	−0.865***	0.680***	–
v_{ua}	−0.993***	−0.841***	−0.892***	0.637**	0.998***
v_{ne}	0.283	−0.166**	0.767**	0.457$^+$	−0.342**
v_{nu}	−0.613	−0.897$^+$	−0.073	0.989***	0.563
v_{na}	−0.913***	−0.642***	−0.986***	0.340	0.937***
v_{w}	−0.998***	−0.927***	−0.793**	0.769***	0.991***
	v_{ua}	v_{ne}	v_{nu}	v_{na}	v_{w}
v_{eu}	−0.993***	0.283	−0.613	−0.913***	−0.998***
v_{en}	−0.841***	−0.166**	−0.897$^+$	−0.642***	−0.927***
v_{ea}	−0.892***	0.767**	−0.073	−0.986***	−0.793**
v_{ue}	0.637**	0.457$^+$	0.989***	0.340	0.769***
v_{un}	0.998***	−0.342**	0.563	0.937***	0.991***
v_{ua}	–	−0.394**	0.515	0.955***	0.983***
v_{ne}	−0.394**	–	0.585	−0.649***	−0.217***
v_{nu}	0.515	0.585	–	0.238	0.665
v_{na}	0.955***	−0.649***	0.238	–	0.883***
v_{w}	0.983***	−0.217***	0.665	0.883***	–

$^+p < 0.10$; $^{**}p < 0.05$ and $^{***}p < 0.01$