Migration Dynamics of Immigrants: Who Leaves, Who Returns and How Quick?

Govert E. Bijwaard*

Econometric Institute
Erasmus University Rotterdam

Econometric Institute Report EI 2005-53

*Erasmus University Econometric Institute, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands; Phone: (+31) 10 40 81424; Fax: (+31) 10 40 89162; E-mail: bijwaard@few.eur.nl This research is financially supported by the Netherlands Organization for Scientific Research (NWO) nr. 451-04-011. I thank Statistics Netherlands and in particular Han Nicolaas and Arno Sprangers, for making the data available to me and for data support. I thank Joop Hartog for valuable comments.

Migration Dynamics of Immigrants: Who Leaves, Who Returns and How Quick?

Abstract

In this paper we analyze the demographic factors that influence the return and repeated migration of immigrants. Using longitudinal data from Statistics Netherlands we track migration histories of recent immigrants to The Netherlands and analyze which migrants will stay in the country, which migrants are more prone to leave and how quick they leave. In order to identify these migrants we apply a mover-stayer duration model on the time spent in the country. We also analyze the return from abroad to The Netherlands of these migrants. Results disclose differences among migrants by migration motive and by country of origin and lend support to our analytical framework. Combining the model for departure from the country and the model for returning to the country provides the long-run stay probability of a specific migrant. It also yields a framework for simulating the life-cycle migration dynamics. The major findings are: (1) labor migrants and students are more prone to leave and migrants who come for family reasons remain in the country more often, (2) migrants from the 'guestworker' countries, Turkey and Morocco, will stay in the country more often than migrants from Western countries.

JEL classification: F22, J10, C41.
Key words: return migration, migration dynamics, mover-stayer model.

1 Introduction

The countries of Western Europe have experienced considerable immigration flows over the past decades and have changed from emigration to immigration countries. In the last decade the majority of immigrants came to Europe for family reasons, family reunification and family formation, while in the 1960s and 1970s the migrants were mainly 'guestworkers', invited low-skilled workers. The most prevalent issue in the recent public debate on migrants in Europe is the assimilation of immigrants in the host country. The assimilation is in general correlated with the time the migrant has spent in the host country. However, migration is not always a once-and-for all decision. Some migrants decide to emigrate again and eventually return, while other migrants leave the host country forever. Despite the knowledge that many migrations are temporary or repetitive the majority of the literature on migration (implicitly) assumes migrations are permanent (notable exceptions are Goldstein (1964), Duleep (1994), Dustmann (1995, 2000, 2002) and Constant and Massey (2003)). In this paper we analyze the characteristics of the immigrants on the decision and timing of leaving and returning to The Netherlands.

The early theories on migration explain the migration flows as a result of wage differentials or through differences in unemployment levels. Given the great and persistent wage and unemployment gaps between most developing countries and the Western World, these conventional migration theories are unable to explain the small size of migration flows and the presence of extensive return-migration. For example, about 20% to 50% of immigrants to the Netherlands leave this country again (CBS 2003). If migration is viewed as an investment decision to maximize human capital and/or earnings over the life-time than return and repetitive migration are not anomalies but common outcomes of a migration decision (see Dustmann (1995, 2000, 2002) and Borjas and Bratsberg (1996)). For example, many migrants have a higher marginal utility of consumption in their home country than abroad (positive location specific externalities due to emotional links) and therefore intend to stay for a short period. However, many migrants change their migration plans during their stay because they are faced with a different situation either in the host or in their home country. Imperfect information about the labor market in the host country is an important reason for return. Changes in the returns to human capital at home or abroad are also motives to return.

Immigrants are a selective sample from the home country population (Chiswick (1978) and Borjas (1994)). Similarly, the immigrants that decide to remain in the host country and the immigrants that decide to leave the host country are a selective subsample of the ones the entered (Borjas and Bratsberg (1996) and Reagan and Olsen (2000)). Subsequently, the decision to return after some time abroad also occurs selectively. Therefore, the immigrants that remain in the country are not a representative sample of the cohort that originally entered, creating potential bias whenever cross-sectional data are used to study patterns and processes of assimilation (Borjas 1985). Regardless of whether emigrants and repetitive migrants are positively or negatively selected, their departure has important implications for the population and economy of a host country. Human capital based theories imply that assimilation in the host country and migration decisions are correlated over time and it is therefore more appropriate to base the analysis of migration on a dynamic model that takes the timing of migration moves into account (Hill (1987) and Dustmann (2002)).

The relevant data for analyzing migration dynamics show three interesting features. First, the timing of the migration events, emigration or immigration, is relevant in understanding the migration dynamics. It is very likely that the assimilation of the migrant in the host country depends on the length of stay in that country. The longer the stay the more opportunities the migrant has had to learn the language and the culture of the host country. Second, along with the migration decisions other relevant characteristics of the individuals may also change over time, like the labor market status and martial status of the migrant. Third, it is hardly ever possible to observe migration decisions over the whole life time of a migrant. These decisions are usually only observed within a predefined period. At the end of this period we only know the current migration status of the migrant. In fact, an immigrant still in the host country may leave that country after this end time. The knowledge that the immigrant has been in the host country from his entry time up till the end, however, contains valuable information. The three data features establish that one has to rely on duration models for repeated events.

Models for duration data were initially developed in the medical sciences and reliability theory (see Oakes (2001)). Duration models or event history models have also been used extensively for demographic analysis, for example in modelling time till birth, time till marriage or time till

death. However, the number of analyses of migration decisions based on a duration model is rather limited and duration analysis of return migration is even more scarce. A few exceptions are Detang-Dessendre and Molho (1999), Longva (2001) and Constant and Zimmermann (2003). Most migration data 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. Therefore, a more common approach is to estimate a probit or logit model for the probability to return (see a.o., Reagan and Olsen (2000) and Constant and Massey (2003)). In a probit model part of the migration dynamics is discarded because only the whereabouts of the migrants at fixed points in time are considered. The choice of these fixed points has a big impact on the estimation results. It is also not straightforward to include time varying covariates into a probit model. In our data the exact moments of immigration and emigration are known. This allows the estimation of duration models.

In conventional duration models it is assumed that in the end all individuals experience the event of interest. That is, all migrants are *movers*. In our application this implies that eventually all immigrants leave The Netherlands. It is, however, very plausible that some of the immigrants never leave. To account for the possibility that some of the immigrants stay forever we use a mover-stayer approach. This approach was developed by Boag (1949) and applied to model the recidivism of criminals (Schmidt and Witte 1989) and labor market transitions (Dunsmuir et al. 1989).[1] To our knowledge a mover-stayer approach has never been used for modelling the migration dynamics. Thus, although in the discussion on migrants it is often assumed that migrants reside permanently in the host country, nobody has attempted to model simultaneously both the intensity to leave and the probability to stay forever.

Upon using a mover-stayer model of the dynamic process of migration decisions, we can, simultaneously, identify the underlying determinants of the timing of this process and the probability to stay forever in the host country or abroad. It will be possible to characterize the immigrants who choose to migrate again, or re-immigrate from a life-cycle perspective. It should be emphasized that the model results will not only indicate who are more prone to move

[1] Schmidt and Witte (1989) use the term 'split-population'model. In the biomedical literature the mover-stayer model is known as cure-model. Maller and Zhou (1996) discuss the implications of such models.

again but also when this move may occur. For a given cohort of immigrants we will be able to simulate a life-cycle of migration dynamics after the first arrival to The Netherlands. Thus, this research can also improve the ability to forecast the immigration population in a country.

The analysis is based on data for The Netherlands. These data contain information on the timing of migration moves to and from the country. A nice feature of these data is that the main migration motive for all non-Dutch immigrants is available. We can, therefore, identify different groups of immigrants and analyze their distinct migration behavior. Labor immigrants, student and marriage-migrants often differ substantially in their migration plans. The country of origin, another important indicator of the migration behavior, is also available. Our results provide strong evidence that some migrants stay forever, while other migrants leave again. We find that the migration motive and country of origin play a very important role in explaining the migration dynamics of the immigrants.

The outline of the paper is as follows. In the next section we present the data and discuss the recent migration pattern to and from The Netherlands. Section 3 discusses the methodology of mover-stayers duration models. Section 4 discusses the empirical results both for the departure from the country as the return to the country. In 5 we use the estimation results to simulate migration dynamics from and to the country over the life-cycle. Section 6 summarizes the results and states our conclusion.

2 Data on immigrants to The Netherlands

In the early 1960s The Netherlands changed from an emigrant to an immigrant country.[2] Immigration follows a European sequence of post World War II and post-colonial immigration, unskilled manpower recruitment and the arrival of refugees. The first period is characterized by the de-colonization of Indonesia in 1949, as a consequence many Indonesian people came to The Netherlands. In the second period, starting in the beginning of the 1960s, a large flow of 'guestworkers', mainly Turks and Moroccans arrived. The Dutch government regulated the recruitment practices by bilateral agreements with the main countries. The total number of

[2]See Zorlu and Hartog (2001) and Van Ours and Veenman (2005) for a more detailed discussion on the immigration to The Netherlands.

immigrants reached 235,000 in 1970. The recruitment policy stopped during the first oil crisis. However, the immigration from the recruitment countries continued as a follow-up migration, first in the form of family reunification and later also family formation. In this period the independence of Surinam also caused large immigration. The final period, starting in the 1980s, is characterized by the family reunification/formation of 'guestworkers'. Additionally, the flow of political refugees, asylum seekers has increased dramatically.

All immigration by non-Dutch citizens who legally entered 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). The CBS, Statistics Netherlands, has linked these data to the Municipal Register of Population (Gemeentelijke Basisadministratie, GBA). These combined data contain information for all migrants on the timing of migration moves and some basic demographic characteristics. The main migration motive for all non-Dutch immigrants is available in the CRV registers, except for those immigrants that leave, die or naturalize before January 1st 1998. The migration motive is also unknown for immigrants that leave before the end of the year of arrival. With these data we can identify important groups of immigrants to the Netherlands. We make the distinction between labor-migrants, asylum seekers (and refugees)[3], family reunification migrants, family-formation migrants, student immigrants and immigrants for other reasons (including a.o. joining with labor migrant, medical treatment and Au Pair). Because we are interested in the migration moves of the potential labor force we restrict our analysis to non-Dutch immigrants between 18 and 64 years of age who arrived after January 1995. The distribution of those immigrants over the migration motives at first arrival to The Netherlands is depicted in Figure 1. Figure 2 shows the development of the absolute numbers of immigrants over the years 1995 till 2003.

– put figure 1 about here–

– put figure 2 about here–

[3]Asylum seekers and refugees are synonymous terms throughout this paper.

From 1995 till 2001 the number of immigrants increased every year. In 2001 69,000 non-Dutch immigrants between 18 and 64 years of age entered The Netherlands. In the last two years the inflow of immigrants decreased to 57,000 in 2003. This decrease is induced by two reasons. First, the more strict asylum policy of the Dutch government has reduced the inflow of refugees from 15,000 in 2000 and 2001 to 5,000 in 2003. Second, the downfall of the Dutch economy has led to a reduction in labor immigrants. In the last 10 years family-formation has been the most important reason to migrate to The Netherlands (26%). Labor migrants (23%) and refugees (17%) are also important groups. Because the migration motive is unknown for the immigrants that leave the country in the same year they entered we have a relatively large number of immigrants with unknown migration motive.

In Table 1 we present some descriptive statistics for the data. Family formation migrants and, of course, students are younger than other migrants. We observe substantial differences in the gender distribution of the migrants. Labor and asylum migrants are mostly men, while migrants for family reasons are mostly women. Not surprisingly, migrants who come for family reasons are more often married, while students are hardly ever married. The table also shows the distribution of the migrants over a selected group of countries/regions of origin. It is immediately clear that the main countries of origin for each migration motive are very distinct. The majority of labor migrants originates from an EU15 or EFTA country. The rest of Asia and USA/Canada are other important countries of origin of labor migrants. The EU15/EFTA region is also an important region of origin of family reunification migrants. Turkey, Morocco, the rest of Asia and the rest of Africa are other important regions these migrants originate from. The same regions are also important regions of origin for family formation migrants. Of course, none of the asylum seekers comes from the EU15 or any other Western country. These migrants mainly come from Africa, Iraq, Afghanistan and former Yugoslavia. The remaining migrants come mostly from the EU and the rest of Asia. Note that the relatively large number of immigrants from Surinam is not reflected in the table because many of them have the Dutch nationality.

Table 2 summarizes the dynamic aspect of migrants. The average observed duration of stay in The Netherlands is the longest for family reunification migrants and the shortest for students. Of all labor migrants that arrive between January 1995 and December 2003, 40% has left the

Table 1: Basic statistics immigrants

	Labour	fam. reunification	fam. formation	Study	Other	Asylum
average age	27	26	24	21	27	26
female	30.8%	70.8%	69.1%	47.0%	67.2%	34.9%
married	12.3%	42.0%	43.0%	2.0%	28.5%	20.5%
			Country of origin			
Europe						
EU15/EFTA	61.5%	15.7%	6.7%	25.7%	43.8%	.
Former Yugo.	0.6%	3.7%	2.5%	1.2%	1.3%	12.8%
new EU	3.9%	2.9%	5.2%	5.4%	6.6%	0.3%
rest of Europe	4.1%	5.0%	6.1%	6.3%	6.0%	9.4%
Asia						
Turkey	2.3%	12.7%	17.8%	1.5%	1.9%	2.6%
China	1.2%	2.0%	2.1%	13.7%	0.8%	0.8%
Iraq	$< 0.1\%$	7.2%	0.6%	$< 0.1\%$	0.4%	17.8%
Iran	0.3%	2.1%	0.8%	0.7%	0.9%	7.1%
Afghanistan	$< 0.1\%$	3.6%	0.6%	$< 0.1\%$	0.2%	14.0%
rest of Asia	8.9%	11.1%	14.1%	19.1%	12.4%	6.5%
Africa						
Morocco	1.0%	11.8%	16.4%	3.4%	1.1%	0.6%
rest of Africa	5.1%	10.4%	9.9%	10.6%	8.5%	27.6%
America						
USA/Can	7.0%	3.3%	3.6%	4.1%	7.7%	.
Surinam	0.2%	3.4%	6.1%	2.4%	1.1%	0.2%
Latin America	2.1%	4.5%	6.9%	5.2%	2.7%	0.2%
Australasia						
Aus/NZ	1.6%	0.8%	0.9%	0.6%	4.3%	.
# observations	114424	37434	129877	46391	35189	95716

Source: Non-Dutch immigrants aged 18-64 to The Netherlands 1995-2003, Statistics Netherlands.

Table 2: Basic duration statistics immigrants

	Labour	fam. reunification	fam. formation	Study	Other	Asylum
duration of stay (mos)	35	58	45	29	38	52
emigration	40%	20%	11%	39%	46%	16%
			Abroad			
duration of stay (mos)	29	32	27	31	33	22
re-immigration[a]	8%	12%	17%	8%	7%	4%
			Repeated migration			
re-emigration[b]	26%	20%	15%	26%	23%	18%

Source: Non-Dutch immigrants aged 18-64 to The Netherlands 1995-2003, Statistics Netherlands.
[a] Percentage of those that have left the country.
[b] Percentage of those that have left the country and have returned.

country by the end of the observation period, December 31st 2003. For migrants who came for family reasons or as a refugee less than 20% has left the country. From the migrants that have left the country asylum seekers hardly, 4%, returns, while the migrants who came for family reasons have a higher return rate.

The simple descriptive statistics in Table 2 give an under-representation of the migration dynamics, because the recent cohorts of immigrants are only followed for a very short period of time. When we only look at the immigrants arriving in 1995 in The Netherlands, we observe that almost 40% of the immigrants from that cohort have left the country within 7 years. From those that have left 16% return to The Netherlands again within the same period (6% of the original cohort). From those returning 33% leave again (2% of the original cohort) in this period. Thus, from these aggregate number we observe already substantial differences among the timing of migration moves for migrants grouped according to their migration motive. This is a strong argument to apply a duration model for the analysis of migration dynamics.

3 A Duration Analysis of Migration Dynamics

Three features of migration-data establish that the most appropriate model for the analysis of migration dynamics is a duration model. In a duration model the timing of a particular event (or recurrent event) is modelled. For many economic and demographic phenomena the timing of a transition from one state into another state is important. Examples are: the time until

death, the time till consecutive births or the time till an unemployed individual finds a job, see a.o. Lancaster (1990) and van den Berg (2001).

The first and very obvious reason to use duration models is that the timing of the migration events, emigration or immigration, is relevant in understanding the migration dynamics. It is very likely that the assimilation of the migrant in the host country depends on the length of stay in that country. The longer the stay the more opportunities the migrant has had to learn the language and the culture of the host country. The second reason to apply duration models is that many relevant characteristics of the migrant may change over time (e.g. martial status or labor market situation). In duration models it is straightforward to incorporate time-varying variables, while in other approaches, like logit or probit, it is not so easy. The third and final reason to use duration models is that it is never possible to observe the complete migration dynamics over the whole life of the migrant, due to a limited observation window. For instance, in our data on newly arriving immigrants to The Netherlands we only observe the in- and outflow of migrants from January 1995 till December 2003. For the migrants still in The Netherlands in December 2003 we do not know their complete length of stay in the country. We only know their migration history up till December 2003. This still contains valuable information and duration models are perfectly fit to use the information of such right censored durations.

The key variables in duration analysis are the duration till the next event, the length of stay of the migrant in The Netherlands, and the indicator of censoring. We denote the duration of individual i for it's j^{th} stay in The Netherlands by T_{ij}. In duration analysis the hazard rate or intensity is usually modelled.[4] In the study of migration dynamics, the intensity gives the instantaneous probability of leaving the country at a duration t months, given that the individual stayed in the country for at least t months

$$\lambda(t) = \lim_{dt \downarrow 0} \frac{P\big[t < T < t + dt | T > t\big]}{dt} = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{S(t)},$$

where $f(t)$ is the probability density function, $F(t)$ is the cumulative density function, and $S(t)$ is the survival function. The intensity is invariant to censoring. The censoring-indicator

[4]In the biomedical literature the accelerated failure time that models the log-duration is also often used. In these models is it more complicated to account for censoring.

$\delta_{ij} = 1$ when the j^{th} stay of migrant i has not been terminated by emigration at the end of the observation window.

A common way to accommodate the presence of observed characteristics is to specify a proportional intensity model

$$\lambda(t|X) = \lambda_0(t)\exp(\beta' X_i(t)), \tag{1}$$

where $\lambda_0(t)$ represents the baseline intensity, that is, the duration dependence of the intensity common to all individuals. The covariates affect the intensity proportionally, see Cox (1972).

3.1 Mover-stayer models

Up to this point we have assumed that all migrants are (potential) *movers*. We now account for the possibility that some migrants never make a next migration move by using a mover-stayer approach. This approach was developed by Boag (1949) and first applied in the social sciences to model the recidivism of criminals by Schmidt and Witte (1989). In labor economics this approach has been used to model 'unemployables' and permanent jobs, see Yamaguchi (1992).

A mover-stayer model assumes that a latent group of individuals have a zero probability to leave, the *stayers*. To incorporate the possibility of defective risks the survival function is redefined as

$$S(t|x_i) = \big(1-p\big)\exp\Big(-\int_0^t \lambda_0(s)e^{\beta' x_i(s)}\,ds\Big) + p, \tag{2}$$

where p is the proportion of stayers. Thus the survival function is given by the proportion of stayers, who never leave the country, plus the proportion of movers multiplied by the probability to migrate after a duration of t months in the country. The proportion of stayers can also depend on observed characteristics of the migrants. To guarantee that the proportion lies between zero and one we employ a logit form: $p(z_i) = 1/\big(1+\exp(\gamma' z_i)\big)$. The intensity is now

$$\lambda(t|x_i) = \frac{\big(1-p(z_i)\big)\lambda_0(t_{ij})e^{\beta' x_i(t_{ij})}\exp\Big(-\int_0^t \lambda_0(s)e^{\beta' x_i(s)}\,ds\Big)}{\big(1-p(z_i)\big)\exp\Big(-\int_0^{t_{ij}} \lambda_0(s)e^{\beta' x_i(s)}\,ds\Big) + p(z_i)} \tag{3}$$

and the contribution to the likelihood of migrant i is

$$\begin{aligned} l_i \;=\; & \prod_{j=1}^{k_i}\Bigg[\big(1-p(z_i)\big)\lambda_0(t_{ij})e^{\beta' X_i(t_{ij})}\exp\Big(-\int_0^{t_{ij}} \lambda_0(s)e^{\beta' X_i(s)}\,ds\Big)\Bigg]^{(1-\delta_{ij})} \\ & \times\Bigg[\big(1-p(z_i)\big)\exp\Big(-\int_0^{t_{ij}} \lambda_0(s)e^{\beta' X_i(s)}\,ds\Big) + p(z_i)\Bigg]^{\delta_{ij}} \end{aligned} \tag{4}$$

where k_i is the number of entries into The Netherlands of migrant i.

3.2 Unobserved Individual Heterogeneity

If there is interdependence of the repeated migrations due to omitted covariates or individual-specific effects, like being adventurous, the parameter estimates may be biased and/or the estimated covariance matrix provides invalid standard errors. One approach is to explicitly model the individual-specific effects using unobserved heterogeneity. In Cox survival models this kind of model is called the mixed proportional hazard model, see for example Manton et al. (1981). The intensity of migrant i at a duration of t months is now given by

$$\lambda(t|v_i, x_i) = v_i \lambda_0(t) \exp(\beta' x_i(t)), \tag{5}$$

where the $v_i > 0$ are i.i.d. random variables with (mixture) distribution function $G(v)$ that differ among the individual but remain the same for each emigration of a particular individual. The Gamma distribution with mean one and variance σ^2 is most often chosen to represent the unobserved heterogeneity. However, if the underlying distribution of the unobserved heterogeneity is not a gamma distribution the results may be biased. Any other mixture distribution, like the normal, or log normal distribution, have the same problem. More robust, and very flexible, is to assume that the mixture distribution can be approximated by a finite discrete mixture (see Heckman and Singer (1984)). For a discrete mixture model, there are a finite number of values or classes, v_l $(l = 1, \ldots, L)$, each having probability q_l $(l = 1, \ldots, L)$ in the population, where $\sum q_l = 1$.

It is important to point out that the presence of stayers is compatible with a discrete mixture duration model. Heckman and Walker (1987) recognize that some specifications of the latent intensity can deliver stayers, like for one particular $l' : v_{l'} = 0$ with $q_{l'} > 0$. The close link between mover-stayer models and a discrete mixture model implies that the two can easily be combined.

We attempt to accommodate the presence of unobserved heterogeneity in the mover-stayer model by assuming that the intensity to leave (for the movers!) comes from a finite discrete mixture. Then, there are two sources of unobserved heterogeneity competing with each other to account for unforseen factors. On the one hand, there is the distinction between movers

and stayers. On the other hand, conditional on being a mover, there are L latent groups of migrants in the intensity. The contribution to the likelihood of migrant i is the likelihood with the distribution of the latent v_l's integrated out

$$
\begin{aligned}
l_i &= \sum_{l=1}^{L} q_l \prod_{j=1}^{k_i} \Big[\big(1-p(z_i)\big) v_l \lambda_0(t_{ij}) e^{\beta' X_i(t_{ij})} \exp\Big(-v_l \int_0^{t_{ij}} \lambda_0(s) e^{\beta' X_i(s)}\, ds\Big) \Big]^{(1-\delta_{ij})} \\
&\qquad \times \Big[\big(1-p(z_i)\big) \exp\Big(-v_l \int_0^{t_{ij}} \lambda_0(s) e^{\beta' X_i(s)}\, ds\Big) + p(z_i) \Big]^{\delta_{ij}} \qquad (6)
\end{aligned}
$$

3.3 Identification issues

A crucial issue is the identification of the latent classes in the discrete mixture mover-stayer model. One of these latent classes is the group of stayers and the remaining classes divide the movers into different groups. We address this issue in a sequential way. First, is it possible to identify the stayers? And second, conditional on an affirmative answer to this question, are the other latent classes identified? Under ideal circumstances, like an infinite observation window, it is obvious that we can distinguish between finite and infinite durations. Thus, if the observation window is sufficiently long we can detect the presence of stayers. The presence of stayers is indicated by an empirical survival function that converges to a positive value. This implies that after a particular duration, the maximum non-censored duration, all remaining durations will always be censored. It is clear that if the observation period is too short there is no hope of identifying the stayers because censored durations can be generated by either movers or stayers.

Maller and Zhou (1996) discuss the non-parametric tests for the presence of stayers and sufficient observation length in the context of a single event (survival till 'death'). In our application of migration dynamics we feel confident an observation window of nine years is sufficient to detect those migrants that stay forever in the country. Detecting permanent stay abroad might be more problematic, because the migrants first should have left the country before the clock starts running abroad and the return from abroad is much slower. Indeed, for the family formation migrants that have left the country, we found a permanent stay probability of zero for the reference migrant. However, we could identify the permanent stay probability abroad for these for some countries of origin.

Heckman and Walker (1987) point out that, in their duration-to-birth example, the popula-

tion of couples without children consist of sterile couples (stayers) and fertile couples that never got children (defecting movers). Unlike stayers, defecting movers have been at risk of moving, but eventually ended up never moving. Such a group exists if the intensity declines sufficiently fast with duration. Crucial in this respect is the behavior of the model in the right tail of the duration distribution. In our application we implicitly assume that none of the movers defect by imposing a, very flexible, piecewise constant baseline intensity. Such a baseline intensity implies that $\lim_{t\to\infty} \int_0^t \lambda_0(s)e^{\beta' x_i(s)}\, ds = \infty$ and, thus, that the defective proportion in the population is zero. Abbring (2002) shows that under this condition the identification question of the other latent classes is answered. In practice, however, these latent classes may be hard to find. The failure of finding any unobserved heterogeneity does not necessarily imply it is not present, see Bijwaard and Ridder (2005), and therefore we should not refrain from introducing a mixture model.

4 Empirical Findings

For our analysis we focus only on the immigrants who enter The Netherlands to work, for family reasons or to study. Asylum seekers are removed from the sample because many of them are not immediately registered in the Municipal Register of Population. Most of the asylum seekers are only registered after they have received a living permission. It can take up to eight years until a living permission is granted. Thus, the registered time in The Netherlands for asylum seekers is smaller than the true duration in the country. Another issue is that some asylum seekers have a temporary permit to stay, awaiting a permanent permit. If the permanent permit is not granted the asylum seeker may be expelled from the country. Then, return migration is an exogenous event which is not based on an individual decision.

The immigrants with other reasons to enter come for a plethora of reasons and therefore they are a very heterogeneous group. The analysis would tell little about the intensity and permanent stay probability of a individual member of this group. Besides, they only amount to 7% of all immigrants. This made us decide not to include these immigrants in the analysis sample and focus on the remaining four groups of immigrants.

Three issues in the data, that can be handled by the mover-stayer model easily, should be

mentioned. First, although in principle the exact date of emigration (and second and repeated immigration) is known, some migrants do not officially inform the municipality where they live that they leave. Their departure is only registered as "administrative removal" after the municipality has assessed that the migrant has left the municipality without showing up in the files of another municipality in The Netherlands or as an emigrant. Administrative removals are included among emigration and they amount to around 40% of the emigration. For asylum seekers the problem of administrative removal is most severe as more than 80% of the emigration of asylum seekers is due to administrative removal. It is quite possible that some migrants that are "administratively removed" remain in The Netherlands as an illegal immigrant. However, an indication that "administrative removal" is not only induced by people who try to stay illegally in the country is that many Dutch nationals also forget to register their move. Those "migrants" very often show up later in the files as an "administrative addition". The non-Dutch, our population of interest, do not very often show up again in the files after an "administrative removal", still we think most of them indeed have left the country. Then an administrative removal implies that the migrant has left before the date the administrative removal is recorded and instead of the true duration we measure the maximum duration of stay for such a migrant. In duration models such information is called *left-censored* data and it can easily be dealt with. For a left-censored duration at t months the contribution to the likelihood is the probability the migrant has stayed at most t months. This is equal to one minus the survival probability up to t months of this migrant.

The second and third data issue both concern the observation of the migration motive. The migration motive is unknown for immigrants that entered the country between January 1995 and December 1997 and left the country before January 1998. This implies that the sample of immigrants that came to The Netherlands in that specific period for whom we observe the migration motive is conditional on 'survival' up to January 1998. In a duration model this is called *left truncation* and by conditioning on 'survival' from the date of entry till January 1st, 1998 we account for this selective observation. The last data issue is similar to the issue just mentioned. The migration motive is also unknown for the immigrants that leave the country before the end of the year. This implies that the immigrants with known migration motive have

'survived' up till the end of the year. Again, conditioning on the time till the end of the year will correct for this selective observation. However, for immigrants that enter the country in the last year of observation, 2003, we cannot apply this. For those immigrants the time till the end of the year equals the observed duration. We, therefore, exclude those immigrants from our analysis.

Other data limitations are that we only have information on a small number of personal characteristics of the migrants. The GBA/CRV registers date of arrival, age, gender, country of origin, immigration motive and marital status. As the construction of the immigrants-panel has just been set-up we thrive to link the data on social-economic characteristics of the migrants in the future.

To formally account for the three issues mentioned we introduce some extra notation. Let a_{ij} indicate whether the j^{th} emigration of migrant i was due to an administrative removal ($a_{ij} = 1$). Let $\delta_{i,98} = 1$ if migrant i arrived before 1998 (and stayed at least till 1998 in The Netherlands). Then $t_{i,98}$ is the duration till January 1^{st}, 1998 for those migrants with $\delta_{i,98} = 1$. Similarly, let $\delta_{i,e} = 1$ for the first entry of migrant i to the country (conditional on arrival in 1998 or later) and let $t_{i,e}$ be the time till the end of the first year of entry. Then the likelihood contribution of migrant i is[5]

$$
\begin{aligned}
l_i \;=\; \prod_{j=1}^{k_i} \Bigg\{ & \Big[\big(1-p(z_i)\big)\lambda_0(t_{ij}) e^{\beta' X_i(t_{ij})} \exp\Big(-\int_0^{t_{ij}} \lambda_0(s) e^{\beta' X_i(s)}\, ds\Big) \Big]^{(1-\delta_{ij})(1-a_{ij})} \\
& \times \Big[\big(1-p(z_i)\big)\Big(1-\exp\Big(-\int_0^{t_{ij}} \lambda_0(s) e^{\beta' X_i(s)}\, ds\Big)\Big) \Big]^{(1-\delta_{ij})a_{ij}} \\
& \times \Big[\big(1-p(z_i)\big)\exp\Big(-\int_0^{t_{ij}} \lambda_0(s) e^{\beta' X_i(s)}\, ds\Big) + p(z_i) \Big]^{\delta_{ij}} \Bigg\} \\
& \times \Big[\big(1-p(z_i)\big)\exp\Big(-\int_0^{t_{i,98}} \lambda_0(s) e^{\beta' X_i(s)}\, ds\Big) + p(z_i) \Big]^{-\delta_{i,98}} \\
& \times \Big[\big(1-p(z_i)\big)\exp\Big(-\int_0^{t_{i,e}} \lambda_0(s) e^{\beta' X_i(s)}\, ds\Big) + p(z_i) \Big]^{-\delta_{i,e}}
\end{aligned}
\tag{7}
$$

The maximum likelihood estimates can be obtained by standard procedures.

[5]The likelihood for a model with (discrete) unobserved heterogeneity is not presented, because in our analyses for none of the migration motives we could identify any latent class.

4.1 Departure from The Netherlands

We estimate four separate mover-stayer models for each of the four migration motives, labor, family reunification, family formation and study. For each migration motive we randomly select 10,000 migrants from our data. Some migrants, 2% to 3%, leave and return again and therefore we have around 10200 recorded stays in The Netherlands for each migration motive. The models consist of three pieces; the baseline intensity, the proportional effect of the covariates on the intensity and the permanent stay probability with the effect of covariates on this probability. First we discuss the baseline intensity.

Many parametric functional forms exist for the baseline intensity. However, they all put heavy restrictions of the shape of the baseline intensity. A more flexible approach is to assume a piecewise constant baseline intensity. Let the intervals $I_m(t) = (t_{m-1} \leq t < t_m)$ for $m = 1, \ldots, M+1$ with $t_0 = 0$ and $t_{M+1} = \infty$ be the intervals on which we define the piecewise constant intensity. Then, the baseline intensity is $\lambda_0(t) = e^{\alpha_0} \cdot \left(\sum_{m=1}^{M+1} e^{\alpha_m} I_m(t)\right)$, with $\alpha_{M+1} = 0$. Thus α_0 determines the intensity in the last interval. The other α's determine the difference in intensity at each interval compared to this last interval. The baseline intensity for a duration of $t \in [t_{m-1}, t_m)$ is higher than the baseline intensity to leave for a duration of $t > t_M$ if $\alpha_m > 0$ and lower if $\alpha_m < 0$.

We have tried a couple of different interval choices and decided to use eight intervals that capture the features of the baseline intensity the best. These intervals are the first 9 months, 9 to 12 months, every half a year till 3 years and 3 to 5 years. We present the results of the estimated (log) alpha's in Table 3. The implied baseline intensities are depicted in Figure 3. We find that the baseline is relatively high with an intensity to leave around 3% to 6% per month after five years in the country. These intensities to leave are, however, conditional on being a mover. As we will see later on, a substantial part of the immigrants will stay forever in the country. When we take this into account the unconditional intensity to leave is much lower and the results are very plausible. The intensity is much higher in the first 9 months.

– put figure 3 about here–

Continuing with the second piece of the model, the influence of the covariates on the intensity

Table 3: Parameter estimates of the baseline intensity to leave

	labor	family reunification	family formation	study
α_0	-3.369^{**}	-3.433^{**}	-2.847^{**}	-3.207^{**}
	(0.134)	(0.151)	(0.216)	(0.199)
α_1 $(0-9)$ mos.	1.171^{**}	1.318^{**}	0.935^{**}	1.199^{**}
	(0.126)	(0.119)	(0.195)	(0.190)
α_2 $(9-12)$ mos.	-0.024	0.067	-0.012	1.381^{**}
	(0.138)	(0.151)	(0.220)	(0.188)
α_3 $(12-18)$ mos.	-0.165	-0.070	-0.351	0.719^{**}
	(0.129)	(0.130)	(0.205)	(0.184)
α_4 $(18-24)$ mos.	0.045	-0.049	-0.178	0.782^{**}
	(0.127)	(0.129)	(0.200)	(0.179)
α_5 $(24-30)$ mos.	0.006	-0.195	-0.364	0.129
	(0.127)	(0.132)	(0.205)	(0.186)
α_6 $(30-36)$ mos.	0.082	-0.167	-0.480^{*}	0.221
	(0.127)	(0.132)	(0.212)	(0.184)
α_7 $(36-60)$ mos.	-0.049	-0.287^{**}	-0.665^{**}	0.117
	(0.111)	(0.107)	(0.177)	(0.161)

Source: Non-Dutch immigrants aged 18-64 to The Netherlands 1995-2003, Statistics Netherlands.
Notes: Standard errors are shown in parentheses. $*p < 0.05; **p < 0.01$.

to leave, it is clear from Table 4 that there are major differences in the regression coefficients estimates across migration motives. Most coefficients refer to the country of origin of the immigrant. We have used the same division of countries/regions of origin as in Table 1 with the European Union/EFTA as reference category. If for a particular migration motive only a few migrants originate from one of these regions the region is dropped from the analysis. The gender, martial status and age (centered at the average age of 25.3 years and divided by 10 and in a quadratic form) are also included.

The possibility that the intensity to leave is different for an immigrant that has been in the country before is captured by the indicator of repeated migration. The assimilation of migrants is probably correlated with the length of stay in the country. The assimilation can accumulate over the periods in the country. However, acquiring information about the host country is costly and migrants who intend to stay only temporarily put less effort in obtaining language and cultural skills. In our data the intended length of stay of immigrants is not observed, but only the real time spent in the country. To allow for less assimilation of intended temporary

immigrants the number of months the migrant has spend in The Netherlands during the previous stay is included in a quadratic form. This implies that 'assimilation' declines in the first months till a 'minimal' length of stay is reached, and than increases with the length of stay.[6]

Some of the regression coefficients have the same sign for the different migration motives. Female and married immigrants always have a lower intensity to leave, although for students neither coefficients are statistically significant. Immigrants from former Yugoslavia have a much lower intensity to leave, up to eight times as low for family reunification (not significant for labor immigrants). Immigrants from Morocco, on the other hand, always exhibit a higher intensity to leave. If the immigrant has been in the country before the intensity to leave is lower than for first time immigrants (not significant for family formation migrants). The coefficients of the length of the previous stay imply a convex effect with a minimum on the intensity for the previous duration. For labor-migrants this minimum is reached around 45 months previously in the country, for family reunification migrants around 32 months and for students around 51 months. This effect is, however, only significant for students. The age effect on the intensity to leave is either convex with a minimum or concave with a maximum. For labor-migrants and family formation migrants the age effect is convex (the latter not significant) with the lowest intensity to leave when the migrant was around 40 years of age. For student-migrants and family reunification migrants the age effect is concave with a maximum intensity around 30 years of age. Labor migrants from Africa, the US/Canada and Latin America exhibit an increased exit rate from the country. Family reunification migrants have a higher intensity to leave when they come from Turkey, China, rest of Asia, USA/Canada, Surinam and Latin America and a lower intensity to leave when they come from Iran. Family formation migrants have a higher intensity to leave when are born in Morocco and a lower intensity to leave when they are born in China, Iraq, Afghanistan or Surinam. Student-migrants have a lower intensity to leave the country when they originate from the new member states of the EU, Turkey, China, rest of Asia, Africa (Morocco excluded), Surinam or Latin America.

For all migration motives there is strong evidence of the existence of stayers, the third part

[6]The length of previous stay is centered per migration motive and divided by 10. The average duration of previous stay used in the centering is 21 months for labor-migrants, 28 months for family reunification migrants, 20 months family formation migrants and, 18 months for students.

of the mover-stayer model. On average $25\% = 1/(1 + \exp(1.115))$ of the labor migrants reside forever in the country. Of migrants that come for family reasons to The Netherlands on average 35% (family reunification) and 51% (family formation) never leave the country again. Not surprisingly, the students exhibit a much lower probability to stay forever of 20%. The impact of the country of origin, gender or martial status on the permanent stay probability differs among the migration motives. Notable exceptions are if the country/region of origin is Morocco, Turkey or the rest of Europe. For all migration motives these countries of origin increase (a negative sign increases the probability) the probability to stay. Female labor migrants have a higher probability to stay forever, while for married labor migrants this probability is lower. The proportion of stayers is higher among labor migrants from the rest of Europe, Turkey, Morocco and the rest of Africa, while this proportion is lower for labor migrants from the rest of Asia and USA/Canada. For family reunification migrants the country of origin has a big impact on the permanent stay probability. On the one hand migrants form countries, like Afghanistan, Iraq, Morocco and Turkey, have a much higher permanent stay probability. On the other hand, a smaller proportion of family reunification-migrants from the USA or Canada reside forever in The Netherlands. For family formation migrants both females and married migrants have a higher probability to stay in the country. Family-formation migrants from Turkey and Morocco are more often stayers, while family-formation migrants from the USA or Canada are less often stayers. Finally, we mention that female students exhibit a higher probability to stay in the country.

Table 4: Parameter estimates of intensity to leave

	labor	family reunification	family formation	study
female	-0.355^{**}	-0.068	-0.511^{**}	-0.022
	(0.068)	(0.081)	(0.099)	(0.055)
married	-0.381^{**}	-0.298^{**}	-0.380^{**}	-0.182
	(0.081)	(0.066)	(0.106)	(0.125)
age	-0.345^{**}	0.055	-0.023	0.299^{**}
	(0.047)	(0.047)	(0.067)	(0.040)
age-squared	0.154^{**}	-0.088^{**}	0.002	-0.183^{**}
	(0.022)	(0.025)	(0.038)	(0.034)
Country of origin				
Europe				
Form. Yugo	-0.755	-2.098^{**}	-1.101^{**}	-1.192^{**}
	(0.397)	(0.199)	(0.328)	(0.198)
new EU	0.158	0.193	0.205	-0.302^{*}
	(0.131)	(0.198)	(0.212)	(0.144)
rest of Europe	-0.040	0.073	0.214	-0.237
Asia	(0.155)	(0.225)	(0.263)	(0.170)
Turkey	-0.512	0.687^{**}	0.177	-1.115^{**}
	(0.340)	(0.135)	(0.181)	(0.186)
China	0.399	0.949^{**}	-1.815^{**}	-1.520^{**}
	(0.264)	(0.128)	(0.349)	(0.083)
Iraq		-0.088	-1.424^{*}	
		(0.453)	(0.623)	
Iran		-1.974^{**}	0.048	
		(0.256)	(0.973)	
Afghanistan			-1.724^{*}	
			(0.782)	
rest of Asia	-0.160	0.276^{**}	-0.007	-0.711^{**}
Africa	(0.093)	(0.105)	(0.165)	(0.066)
Morocco	0.661^{**}	0.597^{**}	0.577^{**}	0.392^{**}
	(0.201)	(0.149)	(0.172)	(0.148)
rest of Africa	0.384^{**}	-0.008	0.196	-0.588^{**}
America	(0.122)	(0.143)	(0.163)	(0.108)
USA/Can	0.437^{**}	0.461^{**}	-0.456	-0.240
	(0.080)	(0.116)	(0.244)	(0.137)
Surinam		1.641^{**}	-1.730^{**}	-2.679^{**}
		(0.114)	(0.212)	(0.223)
Latin Am.	0.595^{**}	0.685^{**}	0.558^{**}	-0.855^{**}
	(0.183)	(0.125)	(0.183)	(0.099)
Australasia	0.082	-0.006	-0.233	
	(0.149)	(0.171)	(0.242)	
Repeated Migration				
repeat mig	-1.380^{**}	-1.521^{**}	-0.249	-1.489^{**}
	(0.175)	(0.300)	(0.308)	(0.189)
t_{-1}	-0.140	-0.036		-0.641^{**}
	(0.112)	(0.120)		(0.182)
t_{-1}^2	0.028	0.043		0.096
	(0.043)	(0.039)		(0.058)

Source: Non-Dutch immigrants aged 18-64 to The Netherlands 1995-2003, Statistics Netherlands.

Notes: Standard errors are shown in parentheses. $*p < 0.05; **p < 0.01$. Base country of origin is EU/EFTA.

Table 5: Parameter estimates for probability to stay

	labor	family reunification	family formation	study
female	−0.537**	0.180*	−0.185*	−0.360**
	(0.070)	(0.078)	(0.087)	(0.080)
married	0.255*	−0.272**	−0.328**	
	(0.129)	(0.069)	(0.089)	
Country of origin				
Europe				
new EU	-0.071	−0.995**	−0.610**	-0.135
	(0.156)	(0.189)	(0.159)	(0.202)
rest of Europe	−0.625**	−1.596**	−1.163**	−1.056**
Asia	(0.147)	(0.162)	(0.174)	(0.142)
Turkey	−0.992**	−1.968**	−1.308**	
	(0.370)	(0.155)	(0.135)	
Iraq		−2.308**		
		(0.174)		
Afghanistan		−3.282**		
		(0.349)		
rest of Asia	0.457**	−0.503**	−0.559**	
Africa	(0.146)	(0.102)	(0.121)	
Morocco	−0.613**	−2.018**	−1.390**	−1.242**
	(0.279)	(0.153)	(0.149)	(0.159)
rest of Africa	−0.593**	−0.772**	−0.662**	0.013
America	(0.127)	(0.106)	(0.132)	(0.151)
USA/Can	1.043**	1.075**	1.079**	-0.160
	(0.123)	(0.167)	(0.324)	(0.164)
Constant	1.115**	0.627**	-0.048	1.386**
	(0.051)	(0.092)	(0.109)	(0.077)

Source: Non-Dutch immigrants aged 18-64 to The Netherlands 1995-2003, Statistics Netherlands.
Notes: Standard errors are shown in parentheses. $*p < 0.05; **p < 0.01$.

4.2 Return to The Netherlands from abroad

About 16% of the migrants arriving in 1995 in The Netherlands leave the country and return before the end of 2003. Thus, even in this relatively short period of nine years, repeated migration is an important phenomenon. As shown in the bottom-part of Table 4, being a repetitive migrant has a large impact on the intensity to leave. For a full coverage of the migration dynamics understanding the chance and speed of return of migrants abroad who first spent some time in the country is imperative. In our data we only observe the migrants while officially registered in The Netherlands. If a migrant leaves the country it is not known whether the migrant returns to its home country or leaves to a third country. Therefore, we cannot include home country characteristics, like the wage differential with the Dutch wages or unemployment level, in the model to explain the return to The Netherlands. However, if the migrant returns we observe the exact time of this new entry to the country and we can link this with the other, already available, information on this migrant. If the migrant has not returned within the observation period, (s)he may stay abroad or return later. This information is rich enough to allow the estimation of mover-stayer models for returning to The Netherlands from abroad. Note again that staying abroad does not necessarily imply that the migrant has resided in one particular country, because we only observe immigration and emigration to and from The Netherlands.

Again we estimate four separate mover-stayer models for returning to the Netherlands. For each migration motive we randomly select 10,000 migrants from the migrants that have left the country. An exception is the family reunification motive. For this migration motive only 7328 migrants have left the country in our observation period and we select all these migrants. We apply again a piecewise constant baseline intensity to return with eight intervals. As before, the regression components of the intensity are gender, martial status, age and most of the countries/regions of origin included in the intensity to leave. For some countries of origin only a few migrants have left, which can be deduced from the estimation results in the previous section, and these countries are dropped from the regression in the intensity to return.

Again to account for assimilation of migrants the number of years the migrant has spend in The Netherlands is included in a quadratic form. This implies that 'assimilation' declines in the first years till a 'minimal' length of stay is reached, and than increases with the length of stay.

It is important to note that the left-truncation issues of the departure from the country are absent in the return from abroad. If we assume that the migrants that have been administratively removed have left the country, administrative removal only induces extra right-censoring for the analysis of the duration abroad. For a migrant that has been administratively removed the exact date (s)he left the country is unknown. This implies that the duration of stay abroad we observe is at least from the administrative removal day till re-entry to the country (or the end of the observation period). This is the same as right censoring. Therefore, we can use the likelihood contribution given in equation (4) instead of the more complicated likelihood contribution given in equation (7). Now $\delta_{ij} = 1$ not only when the migrant is still abroad by the end of the observation period, but also when (s)he has been administratively removed.

We present the estimation results for the components of the baseline intensity in Table 6. The implied baseline intensities are depicted in Figure 4. These baseline intensities are, again, conditional on being a mover. For the migrants abroad a mover will eventually return to The Netherlands, while a stayer never returns (although (s)he might move from one country to another). The return intensity is much lower than the intensity to leave, 0.5% to 1.1% per month after 5 years abroad. Family reunification migrants have a much higher return intensity in the first five years abroad.

– put figure 4 about here–

The differences in the regression coefficients of the intensity to return across migration motives are smaller than the differences in the intensity to leave. Still, there is some notable variety. In Table 7 we present the estimation results. Again the region/country of origin is dropped from the covariates if only a few migrants come from the region/country for the specific migration motive. The number of different countries origin used for the intensity to return is smaller than in the intensity to leave. This is induced by either a high probability to stay in the country (see Table 5) or a low intensity to leave. They both imply that a small number of migrants may leave the country in the time period considered. Migrants form former Yugoslavia (only migration for family reasons), the new EU member states, Turkey and Morocco (not included for labor migrants) have for all migration motives a higher intensity to return from abroad, while migrants

Table 6: Parameter estimates of the baseline intensity to return from abroad

	labor	family reunification	family formation	study
α_0	-5.238^{**}	-4.653^{**}	-5.185^{**}	-4.453^{**}
	(0.957)	(0.631)	(0.212)	(0.459)
α_1 $(0-9)$ mos.	0.720	1.645^{**}	0.846^{**}	0.287
	(0.376)	(0.482)	(0.196)	(0.353)
α_2 $(9-12)$ mos.	0.402	1.309^{**}	0.363	-0.232
	(0.392)	(0.488)	(0.216)	(0.368)
α_3 $(12-18)$ mos.	0.489	1.196^{*}	0.290	-0.227
	(0.371)	(0.479)	(0.207)	(0.347)
α_4 $(18-24)$ mos.	0.560	0.988^{*}	0.087	-0.044
	(0.367)	(0.481)	(0.214)	(0.341)
α_5 $(24-30)$ mos.	0.334	0.921	0.018	-0.595
	(0.373)	(0.483)	(0.221)	(0.360)
α_6 $(30-36)$ mos.	0.045	0.957^{*}	-0.271	-0.562
	(0.388)	(0.485)	(0.237)	(0.365)
α_7 $(36-60)$ mos.	0.033	0.801	-0.222	-0.734^{**}
	(0.354)	(0.468)	(0.208)	(0.333)

Source: Non-Dutch immigrants aged 18-64 to The Netherlands 1995-2003 who have left the country, Statistics Netherlands.
Notes: Standard errors are shown in parentheses. $*p < 0.05; **p < 0.01$.

from the USA and Canada have a lower intensity to return. For some countries of origin the impact on the intensity to return has opposite signs for alternative migration motives. Labor migrants from the rest of Asia return faster to the country while students from the rest of Asia return slower.

The gender of the migrant has only a significant influence on the return intensity of the family reunification migrants, while the martial status is only significant for family formation migrants. Female family reunification return slower, while married family formation migrants return faster after having left the country. The age of the migrant at first entry is an important factor explaining the speed of return. For all migration motives this speed is convex in age with a minimum at around 50 years for labor migrants and around 40 years for the other migrants. The speed of return is also convex in the length of stay in The Netherlands. The speed of returning is decreasing till a duration of about four years in The Netherlands (six years for students) and then increases with the length of stay. A tentative conclusion from this is that, on average, a short stay in the country is ended with a negative selection. With information on the socio-economic status of the migrant this hypothesis could be tested. This information is at this moment lacking in the data.

For three out of four migration motives there is strong evidence of the existence of migrants that stay abroad forever. For the family formation migrants we could only identify a positive probability to remain abroad for a limited number of countries of origin. The relatively short time period of nine years might have caused this. In due time data on more years will become available and it is interesting to check whether this full return to The Netherlands remains. Of the labor migrants who have left the country, on average, $36\% = 1/(1+\exp(0.567))$ stay abroad. This base probability to stay abroad is much higher for family reunification migrants (76%) and students (84%). For family formation migrants the average migrant always returns, but a migrant from former Yugoslavia remains abroad with probability $73\% = 1/(1+\exp(-1.002))$, from the rest of Europe with probability 46%, from Turkey with probability 68% and, from Morocco with probability 46%. Neither gender nor martial status has a significant impact on the probability to stay abroad. An exception is that married family reunification migrants have a lower probability to stay abroad. For those migrants it might be that their spouse is still in The Netherlands when they are abroad.

Table 7: Parameter estimates of intensity to return from abroad

	labor	family reunification	family formation	study
female	0.912	-1.388^{**}	0.055	-0.733
	(0.567)	(0.302)	(0.074)	(0.408)
married	0.416	0.343	0.604^{**}	0.839
	(0.501)	(0.367)	(0.065)	(0.510)
age	-0.276^{**}	-0.392^{**}	-0.335^{**}	-0.680^{**}
	(0.081)	(0.068)	(0.047)	(0.105)
age-squared	0.052	0.112^{**}	0.129^{**}	0.263^{**}
	(0.048)	(0.035)	(0.030)	(0.067)
Country of origin				
Europe				
Form. Yugoslavia		1.209^{**}	2.189^{**}	
		(0.247)	(0.362)	
new EU	0.564^{*}	0.472	0.612^{**}	0.544^{**}
	(0.247)	(0.241)	(0.115)	(0.216)
rest of Europe	0.427	0.410	1.091^{**}	1.217^{**}
Asia	(0.222)	(0.220)	(0.386)	(0.209)
Turkey		0.321	1.235^{**}	0.727
		(0.316)	(0.267)	(0.464)
China	-0.410	-0.402	0.317	1.003^{**}
	(0.592)	(0.381)	(0.253)	(0.220)
Iraq		0.309		
		(0.275)		
rest of Asia	-0.445^{*}	-0.023	0.185	0.858^{**}
Africa	(0.187)	(0.135)	(0.102)	(0.154)
Morocco		0.564^{**}	0.987^{**}	1.333^{**}
		(0.210)	(0.357)	(0.338)
rest of Africa	-0.436	0.734^{**}	0.169	0.477^{**}
America	(0.292)	(0.256)	(0.114)	(0.198)
USA/Canada	-0.666^{**}	-0.591^{**}	-0.516^{**}	-0.397
	(0.204)	(0.213)	(0.155)	(0.309)
Surinam		0.432	0.241	
		(0.272)	(0.147)	
Latin Am.	-0.390	0.759^{**}	0.403^{**}	0.747^{**}
	(0.428)	(0.163)	(0.110)	(0.235)
years in NL	-1.053^{**}	-0.820^{**}	-0.874^{**}	-0.479^{**}
	(0.096)	(0.075)	(0.057)	(0.132)
years in NL, squared	0.141^{**}	0.108^{**}	0.105^{**}	0.038
	(0.015)	(0.011)	(0.009)	(0.024)

Source: Non-Dutch immigrants aged 18-64 to The Netherlands 1995-2003 who have left the country, Statistics Netherlands.
Notes: Standard errors are shown in parentheses. $*p < 0.05; **p < 0.01$.

Table 8: Parameter estimates of probability to stay abroad

	labor	family reunification	family formation[a]	study
female	-1.102	1.581		0.897
	(1.629)	(0.875)		(0.499)
married	1.075	1.061*		0.583
	(0.981)	(0.427)		(0.536)
Country of origin				
Form. Yugoslavia			−1.002**	
			(0.327)	
rest of Europe			0.142	
			(0.606)	
Turkey		-0.370	−0.772**	
		(0.488)	(0.263)	
Morocco			-0.537	
			(0.414)	
rest of Africa		−0.854*		
		(0.384)		
Constant	0.567	−1.130**		−1.642**
	(1.937)	(0.357)		(0.203)

[a] For migrants who came to form a family the probability to stay abroad is only positive for the migrants from the included countries of origin.

Source: Non-Dutch immigrants aged 18-64 to The Netherlands 1995-2003 who have left the country, Statistics Netherlands.

Notes: Standard errors are shown in parentheses. $*p < 0.05; **p < 0.01$.

5 Simulating Lifetime Migration Dynamics

In the previous section we have discussed the estimation of mover-stayer models for both the departure from The Netherlands as the return from abroad back into the country. Some of the migrants will move back and forth and some will stay forever in The Netherlands or abroad (after some time spend in The Netherlands). Migrants that repeatedly move back and forth can, each time they enter, decide to stay forever in the country. Thus, the estimated probability to stay permanently in The Netherlands underestimates the true proportion of migrants that reside forever. This proportion can be derived from the limit proportions of the implied Markov Chain of the combination of both mover-stayer models. This Markov Chain consists of with two states, the migrant is in The Netherlands or the migrant is abroad. If the probability to stay permanently in The Netherlands is denoted by $P_{i,\mathrm{NL}}$ and the probability to stay permanently abroad is denoted by $P_{i,\mathrm{ab}}$, both conditional on the observed characteristics of the migrant available in our data, the long run probability to reside in The Netherlands is

$$\pi_{i,NL} = \frac{P_{i,\mathrm{NL}}}{P_{i,\mathrm{NL}} + P_{i,\mathrm{ab}} - P_{i,\mathrm{NL}} \cdot P_{i,\mathrm{ab}}}. \tag{8}$$

The implied log-run probabilities to reside in the country are shown in Table 9. Most of the family formation migrants reside in The Netherlands. This is induced by the zero probability to stay abroad for these migrants. On average 48% of the labor migrants, 41% of the family reunification migrants and 23% of the students reside, with maybe some time abroad, in the country. Migrants from the USA and Canada have a much lower probability to reside (19% to 26%) in the country, while migrants from Turkey and Morocco mostly (63% to 87%) reside in the country.

Now we have for any specific cohort of migrants entering The Netherlands the long run proportion of residing in the country. The importance of repeated migration can also be derived directly from the permanent stay probabilities $P_{i,\mathrm{NL}}$ and $P_{i,\mathrm{ab}}$. In the long-run the probability to return to the country at least once is equal to $\big(1-P_{i,\mathrm{NL}}\big) \times \big(1-P_{i,\mathrm{ab}}\big)$. Consider, for example, a cohort of unmarried male labor migrants from Turkey entering The Netherlands. From this cohort 71% reside in the long-run in the country (see Table 9). Before they settle down forever, these migrants may have gone abroad temporarily. At first entry 47% of these Turkish migrants

Table 9: Long-run probability to reside in The Netherlands

	labor	family reunification	family formation	study
base	48%	41%	100%	23%
female	47%	53%	100%	35%
married	61%	58%	100%	25%
Country of origin				
form. Yugoslavia			59%	
new EU	49%	66%	100%	25%
rest of Europe	63%	78%	88%	46%
Turkey	71%	82%	85%	
Iraq		88%	100%	
Afghanistan		95%	100%	
rest of Asia	36%	54%	100%	
Morocco	63%	84%	87%	51%
Africa	62%	57%	100%	23%
USA/Canada	24%	19%	100%	26%

Source: Non-Dutch immigrants aged 18-64 to The Netherlands 1995-2003, Statistics Netherlands.
Notes: The base migrant is an unmarried male from a country in the EU or EFTA.

decide to stay forever in the country. The remaining 53% leave the country. From those that have left 36% never return to the country. Thus, $34\% = 53\% \times 64\%$ of the original cohort eventually enters the country again.

– put figure 5 about here–

The time it takes to reach the long-run proportion of permanent stay depends on the intensity to leave the country and on the intensity to return from abroad. With the estimated probabilities to stay and the estimated intensities to leave and return we can simulate, for any cohort of migrants, the migration dynamics after the first entry to The Netherlands. We can consider separate simulations for each possible combinations of migration motive gender, martial status and country of origin, but we only show the simulation for the reference migrant.[7] The reference migrant is, for each migration motive, an unmarried male aged 25 from an EU15 or EFTA country.

[7]As the age of the migrant does neither influence the probability to stay in the country nor the probability to remain abroad, the long-run probability to stay in the country is unaffected by the age. The age does, however, influence the speed of convergence.

For the simulation we consider a cohort of 1000 migrants entering the country together. First we simulate, using the estimated coefficients for the permanent stay probability given in Table 5, which migrants stay forever in the country. For the movers we simulate, using the estimated coefficients for the intensity to leave given in Table 3 and Table 4, for each month the migrants that leave. When a (simulated) migrant leaves we simulate, using the estimated coefficients for the permanent stay probability abroad given in Table 8, whether (s)he remains abroad forever. If the migrant is prone to return we simulate the return on a monthly basis, based on the estimated coefficients in Table 6 and Table 7. If a (simulated) migrant returns to the country we simulate whether (s)he remains forever in the country, etc. We stop simulating 25 years (300 months) after the first entry. We repeat 100 of such simulations and take the average percentage of the original cohort in the country at each month after entry. Figure 5 depicts the evolution of the percentage of the reference migrants still in the country. Note that the long-run permanent stay probability has not been reached within 25 years after entry for the labor-migrants and family formation migrants. For those migrants the percentage of the original cohort still in the country is, after an initial drop, slowly increasing.

6 Conclusion

Most previous studies treat migration as a once-and-for all event and the studies that examine the return and repeated migration neglect the fact that a some migrants reside in the host country forever. For a dynamic analysis of migration from a life-cycle perspective both movers, migrants who leave the country and may return, and stayers, migrants that reside forever in the country, are important. In this paper we addressed the question of return and repeated migration and who resides forever in the host country.

By using a mover-stayer model of the dynamic process of migration we can identify the underlying determinants of the timing of this process and the probability to stay forever in the host country or abroad. If a mover-stayer model for immigrants in the country is combined with a mover-stayer model for migrants abroad, who once stayed in the country, the life-cycle dynamics for a cohort of migrants can be deduced. Some of these migrants decide at their first entry to reside forever in the host country, some other migrants leave after some time in the

host country. Of those migrants that have left the country, some stay abroad forever and some return to the same host country. In the long-run the migrant resides in the country or abroad. This mimics the true dynamic process of migration observed in practice very well.

Using data from Statistics Netherlands for the period 1995-2003 we estimated mover-stayer models both for leaving The Netherlands as for returning to The Netherlands. The data contain information on the timing of migration moves and some demographic characteristics of all migrants entering or leaving The Netherlands in this period. The main migration motive for all non-Dutch immigrants is also available. We can, therefore, identify important groups of immigrants to the Netherlands. For our analysis we use the data on labor-migrants, family reunification migrants, family-formation migrants and student immigrants and select the non-Dutch immigrants who are 18 and 64 years of age at first entry to the country. For each migration motive a separate analysis is performed.

Our results provide compelling evidence that some immigrants to The Netherlands stay forever and some are prone to leave again. We observe big differences in the probability to stay forever and the speed of departure by migration motive and by country of origin of the migrant. Not surprisingly, students are the most prone to leave (and hardly return from abroad) and family formation migrants are the least prone to leave (and return very frequently from abroad). Migrants from countries that used to send guestworkers in the 60s and 70s of the previous century to The Netherlands, in particular Turkey and Morocco, will stay more often permanently in the country than migrants form Western countries.

The statistical advantages of the mover-stayer model is worth emphasizing. First, if movers are present it is obvious that traditional approaches will lead to inconsistent estimates of the regression coefficients in the intensity. Second, neglecting the movers will bias the shape of the baseline intensity towards negative duration dependence because the relative proportion of stayers increase with the duration. This is similar to neglecting unobserved heterogeneity in a traditional duration model. Thus, the concept of movers and unobserved heterogeneity is closely linked.

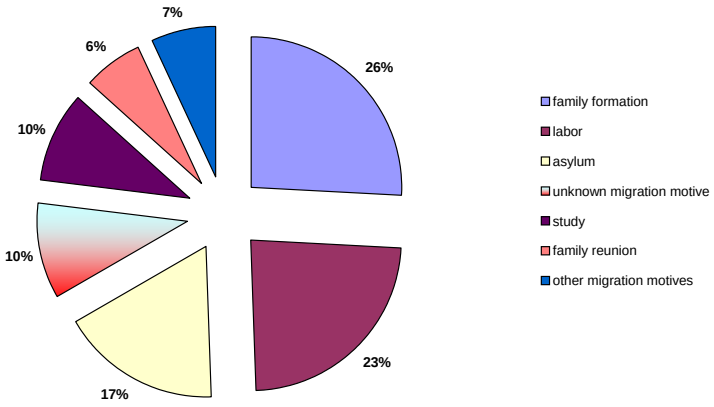

Figure 1: Non-Dutch immigrants (18-64) by migration motive, 1995-2003

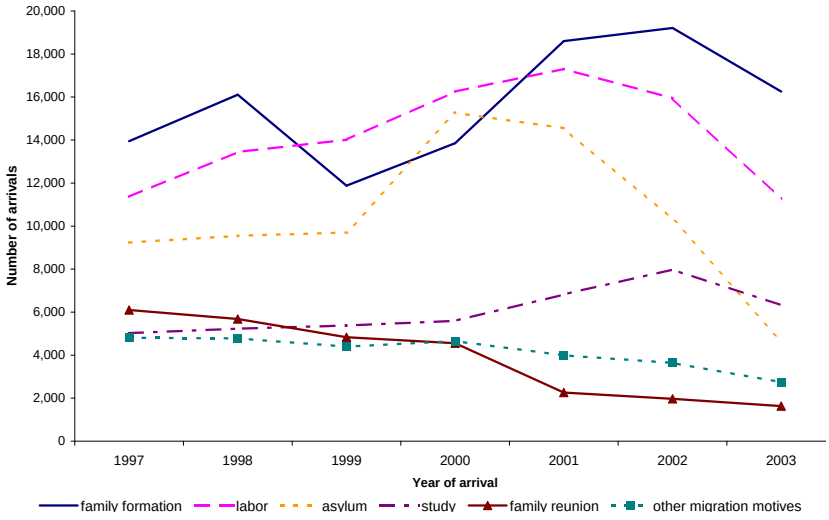

Figure 2: Development of Non-Dutch immigrants (18-64) by migration motive, 1995-2003

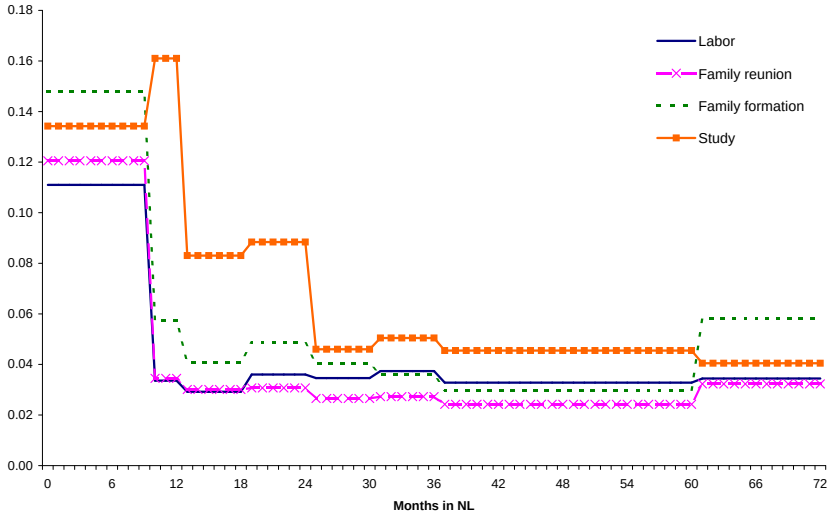

Figure 3: Baseline intensity to leave The Netherlands

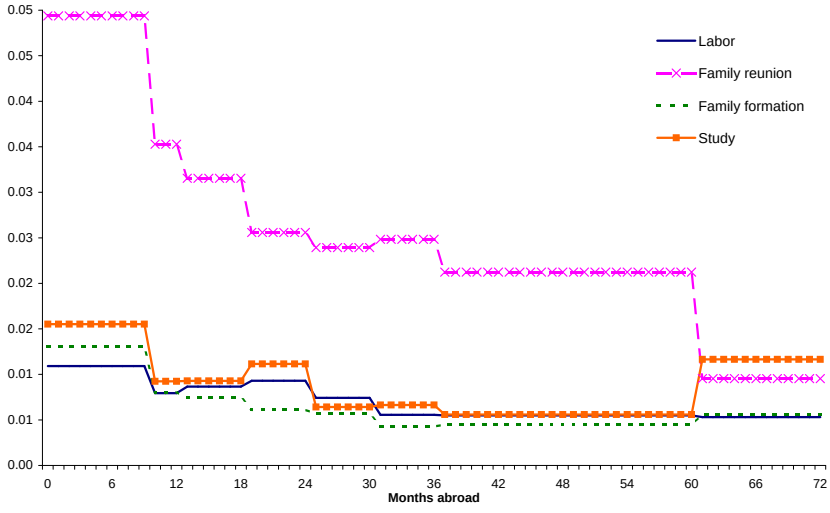

Figure 4: Baseline intensity to return form abroad to The Netherlands

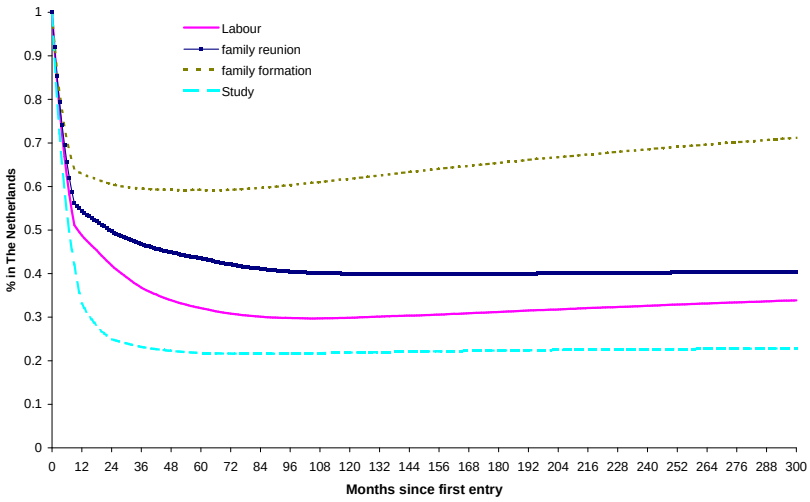

Figure 5: Simulated migration dynamics for a cohort of 25 year old unmarried males from EU/EFTA

References

Abbring, J. H. (2002). Stayers versus defecting movers: A note on the identification of defective duration models. *Economic Letters 74*, 327–331.

Bijwaard, G. E. and G. Ridder (2005). Correcting for selective compliance in a re–employment bonus experiment. *Journal of Econometrics 125*, 77–111.

Boag, J. W. (1949). Maximum likelihood estimation of the proportion of patients cured by cancer therapy. *Journal of the Royal Statistical Society: Series B 11*, 15–44.

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

Borjas, G. J. (1994). The economics of immigration. *Journal of Economic Literature 32*, 1667–1717.

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

CBS (2003). *Allochtonen in Nederland 2003*. Voorburg.

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.

Constant, A. and K. F. Zimmermann (2003). Circular movements and time away from the host country. Discussion Paper 960, IZA, Bonn.

Cox, D. R. (1972). Regression models and life–tables (with discussion). *Journal of the Royal Statistical Society: Series B 34*, 187–220.

Detang-Dessendre, C. and I. Molho (1999). Migration and changing employment status: A hazard function analysis. *Journal of Regional Science 39*, 103–123.

Duleep, H. O. (1994). Social security and the emigration of immigrants. *Social Security Bulletin 57*, 37–52.

Dunsmuir, W., R. Tweedie, L. Flack, and K. Mengersen (1989). Modelling of transitions between employment states for young Australians. *Australian Journal of Statistics 31A*, 165–196.

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

Dustmann, C. (2000). Temporary migration and economic assimilation. *Swedish Economic Policy Review 7*, 213–244.

Dustmann, C. (2002). Return migration, wage differentials, and the optimal migration duration. *European Economic Review 47*, 353–369.

Goldstein, S. (1964). The extend of repeated migration: An analysis based on the Danish population register. *Journal of the American Statistical Association 59*, 1121–1132.

Heckman, J. J. and B. Singer (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica 52*, 271–320.

Heckman, J. J. and J. R. Walker (1987). Using goodness of fit and other criteria to choose among competing duration models: A case study of Hutterite data. In C. C. Clogg (Ed.), *Sociological Methodology (1987)*, pp. 247–308. American Sociological Association: Washington DC.

Hill, J. K. (1987). Immigrant decisions concerning durations of stay and migratory frequency. *Journal of Development Economics 25*, 221–234.

Lancaster, T. (1990). *The Econometric Analysis of Transition Data*. Cambridge University Press.

Longva, P. (2001). Out–migration of immigrants: Implications for assimilation analysis. Memorandum 04/2001, University of Oslo.

Maller, R. A. and X. Zhou (1996). *Survival Analysis with Long-term Survivors*. Chichester: John Wiley.

Manton, K. G., E. Stallard, and J. W. Vaupel (1981). Methods for the mortality experience of heterogeous populations. *Demography 18*, 389–410.

Oakes, D. A. (2001). Biometrika centenary: Survival analysis. *Biometrika 88*, 99–142.

Reagan, P. B. and R. J. Olsen (2000). You can go home again: Evidence from longitudinal data. *Demography 37*, 339–350.

Roodenburg, H., R. Euwals, and H. ter Rele (2003). *Immigration and the Dutch Economy*. The Hague, The Netherlands: CPB Netherlands Bureau for Economic Policy.

Schmidt, P. and A. D. Witte (1989). Predicting criminal recidivism using 'split population' survival time models. *Journal of Econometrics 40*, 141–159.

van den Berg, G. J. (2001). Duration models: Specification, identification, and multiple duration. In J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics, Volume V*, Chapter 55, pp. 3381–3460. Amsterdam: North–Holland.

Van Ours, J. and J. Veenman (2005). The netherlands; old emigrants - young immigrant country. In K. Zimmermann (Ed.), *European Migration: What do we know?*, pp. 173–196. Oxford UP.

Yamaguchi, K. (1992). Accelerated failure-time regression models with a regression model of surviving fraction: An application to the analysis of "Permanent Employment " in Japan. *Journal of the American Statistical Association 87*, 284–292.

Zorlu, A. and J. Hartog (2001). Migration and immigrants: The case of The Netherlands. Discussion paper 2001-042, Tinbergen Institute.