Life Paths of Migrants: A Sequence Analysis of Polish Labor Migrants’ Family-Life Trajectories

Tom Kleinepier
_Netherlands Interdisciplinary Demographic Institute_

Helga A.G. de Valk
_Netherlands Interdisciplinary Demographic Institute and Vrije Universiteit Brussels_

Ruben van Gaalen
_Statistics Netherlands_

_DRAFT - Please do not cite without permission of the author_

All correspondence to:
Tom Kleinepier
Netherlands Interdisciplinary Demographic Institute (NIDI)
Lange Houtstraat 19
2511 CV The Hague, The Netherlands
phone: +31 70 35 65 200
e-mail: kleinepier@nidi.nl
Abstract. In the past decade intra-European migration has become more and more important. The Polish migration is an example of this large and growing group of immigrants in Europe including the Netherlands, for which not much is known yet. Using unique panel data from the Dutch population registers, we apply sequence analysis to explore and describe the variety of family life-courses among young adult (aged 20-30 at migration) Polish labor migrants in the Netherlands. Each individual’s family-life career is followed for five years on a monthly base. Optimal matching analysis is applied to compute (dis)similarities between individuals, in order to group them into predominant life paths. A series of logistic regression analyses on the acquired typology shows a strong impact of partner origin for the type of union and likelihood of return migration. Furthermore, men are more likely to remain single and to leave the Netherlands, while women more often cohabit.

INTRODUCTION

The family life dynamics of international migrants have become a core topic in many demographic studies over the past decades. Despite this great deal of research, less attention has been given to investigating the possible relationship between the family life course of migrants and the migration move itself (Milewski 2007). However, family life experiences and migration events strongly interact with each other. Previous studies suggest that migration increases the likelihood of union dissolution (Frank & Wildsmith 2005) and may alter fertility behavior (Milewski 2007, Lindstrom & Giorguli Saucedo 2002). At the same time, having a partner and/or children decreases the likelihood of long distance (Mulder & Wagner 1993) and return migration (Dustmann 2003). The timing and sequencing of the migration move in the life course may thus be crucial for choices made in the family domain and vice versa.

A common approach in life course research is to apply event history analysis and focus on the occurrence and the timing of a specific event, or combinations of events (e.g., leaving the parental home, union formation, becoming a parent). While event history allows one to study causal relationships and complex interdependencies between specific life course events, it has “no conception of the career as a whole” (Abbott 1990:140). By focusing on events, one may miss the more holistic insight into the life course. This is unfortunate as different events in the life course are not separate experiences but are strongly interdependent (Giele & Elder 1998). In this study, we therefore apply a methodology of research that has become known as sequence analysis (see Abbott & Tsay 2000). This technique allows one to study whole trajectories as a unit of analysis and thus provides a more integral picture of migrants’ family life courses.
We focus on young adult Polish migrants that have recently arrived in the Netherlands. Polish migration to the Netherlands has increased sharply in the past decade and Polish nationals are the largest migrant group settling in the country nowadays (Statistics Netherlands 2013). The majority of recent Polish migrants in the Netherlands has migrated at the age of making the transition to adulthood. Hence, they are a highly interesting group for studying the interaction between migration and family life. Nevertheless, much research on the family dynamics of immigrants in the Netherlands still focuses on the traditional origin groups (Kalmijn & van Tubergen 2006, Zorlu & Mulder 2011, Garssen & Nicolaas 2008), while knowledge on those of Polish origin in the country remains scarce (see Dagevos 2011, Korf 2009, Schothorst 2009 for exceptions). Although there are some recent studies on Polish migration to other European countries, such as the UK (White 2011, Drinkwater et al. 2009, see also the papers in Burrell 2009) and Germany (Nowicka 2013), these papers have mainly been focused on labor market outcomes and neither of these studies has focused on family demographic transitions of Polish migrants.

Polish migration to the Netherlands is often temporary: almost 60 percent of Polish migrants who have come to the Netherlands in the past decade have left the country within seven years (Nicolaas 2011). In this regard, Polish migrants differ strongly from migrants from outside the European Union (EU) for whom it is much more difficult to travel back and forth between the Netherlands and the country of origin due to compulsory visa requirements (Dagevos 2011). Polish nationals can in principle come and go as they please after the opening of international borders between different member states of the EU (Nicolaas 2011). The focus of this study is therefore not only on the association between migration to the Netherlands and the family life course of Polish migrants, but also on how their life paths after migration are related to return and circular migration. We include (return) migration as one of the events in the life course and link this to several events in the family domain: cohabitation, marriage, childbirth, and separation/divorce. The paper is guided by the question: To what extent does the timing and sequencing of migration affect the family life trajectories of Polish migrants in the Netherlands and how are these related to return/circular migration?

The present study contributes to the literature in three ways. First, by applying sequence analysis, we cover multiple family life events in the transition to adulthood simultaneously and provide a holistic picture of migrants’ family life courses. Second, we study the dynamic interplay between family life events and return/circular migration. Much of the theoretical and empirical literature considers migration to be permanent even though temporary migration outnumbers long-
term migration in many regions of the world (Appleyard 1988, Constant & Zimmermann 2011). Finally, we do so by focusing on the rather understudied, yet important, Polish migrant group in the Netherlands.

To address our research question, we apply sequence analysis on rich longitudinal register data from the Social Statistical Database (SSD) that cover the entire population of recent immigrants to the Netherlands. The data contain detailed information on the timing of migration and family life events, allowing to reconstruct the life course trajectory for every individual migrant. Our analyses focus on young adult Polish labor migrants (aged 20-30 at immigration) who have recently arrived in the Netherlands. We thus cover a very dynamic period of the life course, in which many demographic life events occur in a condensed period of time (Rindfuss 1991). The life course trajectories of these migrants are followed for a period of five years on a monthly time base.

BACKGROUND

Polish migrants in the Netherlands

Various groups of immigrants have come to the Netherlands since the middle of the 20th century. These immigration flows were initially dominated by migrants from (former) colonies (e.g., Surinam and the Netherlands Antilles) and labor migrants from Mediterranean countries (e.g., Turkey and Morocco), followed by family reunification in the 1970s and 1980s. The number of asylum seekers increased significantly in the 1990s, but this group still constitutes a relative small proportion of immigrants with a wide variety of backgrounds. In the past decade, a large share of immigrants to the Netherlands came from countries in Eastern Europe, in particular Poland (Statistics Netherlands 2013). Although Polish migration to the Netherlands is not solely a recent phenomenon, the influx of Polish migrants to the country has only increased sharply since Poland joined the EU in May 2004 (Dagevos 2011). This entitled Poles to settle freely in EU countries, although in the Netherlands an employment permit was still required for labor migrants until May 2007 (Korf 2009). In this study, we exclusively focus on Polish labor migrants who have come to the Netherlands after Poland’s accession to the EU, but before the obligation of an employment permit was abolished.

Since 2004, Polish nationals constitute the largest group of immigrants arriving in the Netherlands. At the same time, a large share also return to Poland (Statistics Netherlands 2013). About 90 percent of Polish migrants who have left the Netherlands after 2003 returns to Poland, which implies that there is no substantial onward migration to other EU countries (Dagevos 2011).
Despite the relative high levels of return migration, the net migration rate has been strongly positive since 2004 and has varied between 5,000 to 11,000 persons per year (Statistics Netherlands 2013). As a result, the number of Polish residents in the Netherlands has increased substantially over the past decade. On January 2012, more than 100,000 Poles were entered in the population registers, compared to about 35,000 in 2004 (Statistics Netherlands 2013). Some 25 percent of the Polish community in the country nowadays is born in the Netherlands, but has at least one parent who is born in Poland (second generation). The second generation is considerably smaller than that of the traditional migrant groups in the Netherlands, which is connected to the fact that a large share of first generation of Polish migrants only recently arrived in the Netherlands.

Important is that these statistics are based on Polish migrants who are entered in the population registers. The registration is based on the immigrants’ settlement intentions: only those who intend to stay in the Netherlands for more than three months have to register themselves in the population database of their municipality of residence. This implies that migrants who come to the Netherlands for a short period (predominantly labor migrants) are not included in the registers (see also Dagevos 2011). According to the most recent estimates (Van der Heijden et al. 2013), the majority (77%) of Polish people living in the Netherlands are registered. The numbers most likely vary over the course of the year, peaking in the summer due to seasonal influences (Dagevos 2011).

Many Poles come to the Netherlands for work: nearly 70 percent of the registered Polish migrants in the Netherlands aged between 15 and 65 is in paid employment. This is comparable to the employment rate of the native Dutch population and more favorable than that of non-Western origin groups in the Netherlands (Dagevos 2011). The completed level of education of Polish migrants is lower than that of the Dutch, but much higher than that of non-Western origin groups (Dagevos 2011). Previous studies have shown that there is a strong mismatch between Polish migrants’ qualifications and the jobs they carry out (often under-qualified) (Schothorst 2009). They are often employed in low-skilled jobs in industry, construction and agriculture on temporary contracts and are thus susceptible to unemployment. The Dutch minimum wage is nevertheless usually higher than the salary they could earn in their own occupation in Poland (Dagevos 2011). Finally, the majority of Polish migrants has a poor command of the Dutch language, which is not surprising given their relative short time of residence in the Netherlands (Schothorst 2009, Dagevos 2011).

*Family events in Poland*

To be completed.
MIGRATION AND FAMILY EVENTS: THEORETICAL BACKGROUND

Migration as predictor of family events
The ways in which migration may affect the family life course has been subject of a host of studies. In particular the fertility behavior of migrants after an (international) migration move has been studied in great depth in both the United States (Ford 1990, Carter 2000, Lindstrom & Giorguli Saucedo 2002) and Europe (Milewski 2007, 2010, Kulu 2005). Other studies have considered the relationship between migration and partnership behavior (Boyle et al. 2008, Landale & Ogena 1995, Frank & Wildsmith 2005). The foundation of these studies is centered around theories of socialization, adaption, and disruption. Many authors emphasize the importance of the migrant’s socialization process. Socialization theory (Glass et al. 1986, Inglehart & Baker 2000) argues that preferences are shaped by dominant norms and values during a person’s childhood and remain rather stable throughout the life course. The socialization hypothesis therefore states that migrants follow the union and fertility patterns of their country of origin, even if it differs from that of the host society. Several studies in Western Europe (Andersson & Scott 2007, Milewski 2010) and the U.S. (Massey 1981, Kahn 1988) find support for this hypothesis regarding fertility behavior. De Valk & Liefbroer (2007) find socialization effects with regard to the type of first union (cohabitation vs. marriage) among immigrants in the Netherlands.

In contrast, the adaption hypothesis emphasizes the importance of the destination after the migration move, rather than the environment during childhood. This hypothesis states that migrants adjust their family life behavior according to the economic and cultural environment of the destination country. Frank & Wildsmith (2005) attribute elevated risks of union dissolution among Mexicans after moving to the U.S. partly to changes in normative values and lower levels of social control. Other evidence for the adaption hypothesis comes from studies that observe a convergence in fertility levels between the migrant and majority population (Milewski 2007, Ram & George, 1990, Mayer & Riphahn 2000). The theoretical foundation of this hypothesis is two-fold. First, economists interpret the number of children as a combination of the household income and the relative costs of children (Becker 1981). Moreover, couples may decide upon divorce/separation more easily if affordable independent housing is available (Mulder 2006). Migrants’ family life behavior is thus influenced by the socio-economic situation in the destination country. Second, the resemblance of the migrants’ family life behavior to that of the majority population may also be triggered by norms and values dominant in the destination country. The latter view stems from classical studies on
assimilation theory (see Gordon 1964, Alba & Nee 1997). Since it is hard to separate the effect of cultural factors from that of socio-economic conditions, many authors combine these views under the label of adaption or assimilation (c.f. Lindstrom & Giorguli Saucedo 2002).

Whereas the two previous discussed theories emphasize the importance of the environment either at origin or destination, the disruption hypothesis focuses on the role of the migration process itself. Migration is a stressful life event and often involves psychical separation of spouses for extended periods of time (Frank & Wildsmith 2005). Migration may therefore lead to reduced fertility after the move (Lindstrom & Giorguli Saucedo 2002, Carter 2000) and a higher risk of union dissolution (Boyle et al. 2006, Landale & Ogena 1995, Frank & Wildsmith 2005). At the same time, it has been argued that migration, marriage, and the start of childbearing are closely connected. In the case of marriage-related migration one would therefore expect elevated rather than depressed fertility levels shortly after the move (shown by Lindstrom 2003, Milewski 2007, Andersson 2004). In addition, studies in the United States suggest that couples may be anxious to have a birth after migration in order to grant U.S. citizenship to the child, which opens the route of a legal migration status for the parents (Lindstrom & Giorguli Saucedo 2007).

Migration as outcome of family events
Migration, as well as return migration, has also been considered to be an outcome of family life events. Studies have analyzed the effect of having a partner and/or children on migration behavior. When summarizing the major findings of these studies, it is important to distinguish between short and long distance (international) moves. The literature suggests that the presence of a (married) partner (Sandefur & Scott 1981, Speare & Goldschneider 1987, Mulder & Wagner 1993) and children (White et al. 1995, Kulu 2007) reduces the wish and likelihood of moving over long distances. This decreased mobility is linked to higher economic and psychological costs of relocating a larger family, in particular when some children are of school age (White et al. 1995). In addition, Mulder & Wagner (1993) find that the higher likelihood to migrate among singles can for a large part be attributed to moves related to marriage (interrelatedness of events).

The relationship between family events and return migration has hardly been explored, even though the level of return migration has been high for both the U.S. and in many European countries in the past decades. However, some of the theoretical mechanisms outlined above may also apply to return migration. With regard to partnership behavior, it is crucial where the partner is born. Those who find a partner in the host country are less likely to return to the country of origin (Bijwaard &
Wang 2013, Zhao 2002). Spousal separation appears to be a dominant factor in causing a migrant to return (Zhao 2002). Finally, Bijwaard & van Doeselaar (2012) observe that divorce is an important trigger of return migration. For the effect of the presence of children on the likelihood of return migration mixed evidence is reported. A study by Lindstrom & Giorguli Saucedo (2007) shows that Mexicans in the U.S. are significantly more likely to return to Mexico briefly after the occurrence of the first birth. After the third birth, however, the risk of return drops considerably. The authors suggest that this may be due to the fact that the oldest child may be of school age by that time (see also White et al. 1995). Dustmann (2003) studies the impact of children on return intentions and realizations among migrants in Germany. The analyses reveal a negative association between the presence of children and return migration, regardless of the number of children.

**Contribution of the current study**

The literature overview shows that migration and the family life course are closely connected to each other. Despite this acknowledgement, most research has applied event history analysis, focusing on the time to a specific event. In this study, we provide a more integral picture of the interrelationship between migration and family events by sequence analysis. This technique identifies and visualizes an entire life course of interlinked states, where the timing, sequencing, and quantum of events can be accounted for (Billari et al. 2006). The main objective of this paper is to explore and describe the variety of distinct life paths that Polish migrants in the Netherlands follow. We focus on how the timing and sequencing of migration in the life course interacts with what choices are made in the family domain and how this is related to return and circular migration.

Our study builds on insights from the life course perspective in order to study the relation between family events and migration. Four fundamental principles characterize the life course approach: human agency, time and place, linked lives (social relations), and the timing of life events (Giele & Elder 1998). The life course approach thus assumes that individuals shape their lives as human agents. At the same time, however, life courses are embedded in and shaped by the time and place in which individuals live. Following the previously discussed socialization theory (Glass et al. 1986) one would expect a strong influence of the cultural context in Poland on family lives of Polish migrants in the Netherlands. The fact that the Polish migrants in our study migrated as young adults, implies that they are primarily socialized in their country of origin. We may therefore expect that Polish migrants follow relatively traditional union and fertility patterns in line with those in their country of origin (Sobotkta 2008, Kotowska et al. 2008). This may for example be reflected in a
relative young age of marriage and low incidence of divorce, non-martial cohabitation and out of wedlock births. However, this theory completely neglects the influence of adult experiences. Polish migrants can also adapt to the Dutch society, due to cultural factors and socio-economic conditions (Milewski 2007).

The principal of linked lives refers to the idea that lives are lived interdependently. This principle thus conveys importance of viewing life courses as coordinated and interdependent. As is shown in previous studies, having a spouse and/or children is indeed strongly related to migration intentions and behavior (Dustmann 2003, Mulder & Wagner 1993). We may expect that when the migrants find their spouse in the Netherlands, they are more likely to adapt to the host society. In other words, those with a Dutch partner are expected to demonstrate less traditional family life behavior (e.g., cohabitation and births outside the wedlock). In line with previous studies, we also expect that having a Dutch partner strongly reduces the risk of return migration (Bijwaard & Wang 2013, Zhao 2002).

Finally, the principal of timing suggests that the timing and sequencing of life course events have a large impact on subsequent behaviour and outcomes. The consequences of migration are thus closely linked with the life stage in which the move occurs. Note that this is in line with the previously discussed disruption hypothesis. Previous studies have shown that couples without children are most likely to migrate together, while the likelihood of a men’s solo migration strongly increases after childbearing (Cerruti & Massey 2001). Finally, Dustmann (2003) notes that those who migrate at a later ages may have more difficulties to adopt to the foreign life style, and are likely to maintain stronger links to the country of origin. Consequently, a higher age at migration increases the risk of return migration (Dustmann 2003).

**DATA AND METHODS**

**Data**

In this paper we make use of rich administrative micro panel data that cover the entire population of the Netherlands: the Social Statistical Database (SSD). The SSD was constructed by Statistics Netherlands in the late nineties, by linking several registers to one another. The backbone of this database is the Municipal Personal Records Database (*Gemeentelijke Basisadministratie*, GBA), as all the other files are matched to this register (Houbiers 2004). All immigrants who intend to stay in the Netherlands for more than 90 days are legally obliged to register themselves at the GBA within five
days of arrival. The registration triggers the start of other processes and a proof of the registration is necessary for many other formalities. The data thus contain highly reliable information on the exact date of the immigrants' migration move to the Netherlands. Note that immigrants who stay for a very short period (< 3 months) in the Netherlands are less well represented in the data.

Those who leave the Netherlands are deregistered from the GBA if the expected stay abroad in the year following the departure is at least eight months. However, not all emigrants inform the authorities of their intention to leave the Netherlands. If the authorities find out that a person has left the municipality without showing up in the files of another municipality in the Netherlands or as an emigrant, they are deregistered as an “administrative removal” (see Alders & Nicolaas 2004). These administrative removals are included among emigration statistics and they account for about 40 percent of all emigrations in our sample. Since the exact date of the departure is unknown, we assume that these migrants have left the Netherlands on the day that they were administratively removed. If a migrant comes back to live in the Netherlands after a while, re-entry in the GBA is possible due to a unique personal registration number. This allows us to reconstruct the migration career for each individual, in which potential circular migration is recorded.

In addition to the date of entry and exit, the data provide information on the migrant’s marital status and household composition on a daily basis. For sake of simplicity, we use a monthly time scale to construct a sequence-type representation of family life course and migration events. The representation is based on the migrant’s union status, childbearing events and migration move(s). First of all, we treat cohabitation and marriage as different states. Individuals who were registered as the only person living at a particular address and those who were registered at the same address with people other than parents or partners are classified as single. Those who migrated with their parents are excluded from the analyses, as we do not perceive this to be an independent move. Second, through the record linkage of parents and children, we are able to distinguish married and cohabiting persons into those with children living in the household and those without. Finally, since migration may involve spousal separation (Frank & Wildsmith 2005, Lindstrom & Giorguli Saucedo 2002), we further distinguish between those who are married and living at the same address as their partner, and those who are married but living without their partner.

It is important to note that this information is only available for the migrant’s time of residence in the Netherlands. The data do not provide information on the household composition prior to the migration move. This means that, for example, those who are classified as living single in the Netherlands may have an unmarried partner and/or children living abroad. Similarly, changes in
the family domain that occur after return migration are unrecorded. However, when a migrant re-enters, the information on the previous stay(s) is available, implying that we can observe possible changes in the family domain in the period between return migration and re-entry in the Netherlands.

We focus on young adult Polish migrants who have arrived in the Netherlands after the expansion of the EU in May 2004. The analyses are restricted to those who were 20-30 years old when they entered the country for the first time. We follow each individual life trajectory for a period of five years after the first migration move, so that all sequences are of equal length. Data were available until January 2012. We thus focus on Polish young adults that have migrated to the Netherlands between May 2004 and January 2007. Finally, we only selected those who came to the Netherlands for work related reasons. The migration motive is recorded by the Immigration and Naturalization Service (IND) and has been linked to the GBA. The reason for this restriction is three-fold. First, the vast majority of Polish migrants to the Netherlands migrates for work related reasons. Second, and more importantly, certain migration motives (e.g., family reunification) are strongly related to subsequent family life behavior. Since we are in particular interested in how the timing and sequencing of migration move(s) interact with transitions in the family domain, we aim to have a relative homogeneous group in terms of migration purposes. Finally, the group of Polish labor migrants includes a substantial share of both men (N=3,389) and women (N=1,966), which is not the case for other migration motives. In total, our panel thus includes 5,355 Polish migrants aged 20-30 at migration.

Methods
We analyze the interplay between the Polish young adults’ migration and family life careers by applying sequence analysis (see Abbott 1995, Abbott & Tsay 2000). This technique was originally developed in biology to analyze strings of DNA, but has also proved to be a fruitful approach in the social sciences for the analysis of longitudinal data (Aassve et al. 2007, Elzinga & Liefbroer 2007, Bras et al. 2010). Most applications in the social sciences study life course processes, such as labor market careers, educational careers, or family formation. In contrast to the often applied techniques of event history analysis in life course research, which focus on the time to a single event or combinations of events, sequence analysis respects the whole trajectory and allows thereby a holistic perspective on the life course (Billari & Piccarreta 2005).
In sequence analysis, each individual life course trajectory is represented as a string of characters (states). In this study, we distinguish between seven possible states: single (S), cohabitating (C), cohabitating and child(ren) (CC), married and living without partner (MS), married and living with partner (MP), married and child(ren) (MC), and return migration (R). Clearly, distinguishing between more states (e.g., the number of children living in the household) may be appealing on theoretical grounds. However, each distinction introduces another sequence state and complicates the analysis. Thus, for simplicity, we only distinguish between whether a couple has children living in the household or not. Consequently, there is no change in parenthood status in the case of second and subsequent births, nor if one of multiple children leaves the parental home. Similarly, we do not distinguish between individuals who are living alone and those who are sharing their residence with people other than partners and parents. It is our opinion that classifying the latter as single is also justified from a theoretical perspective, since the commitment to these other household members is probably low (see Korf 2009). Finally, we do not treat singles with children and singles without children as different states, as the number of sequences including single parenthood was found to be very limited (N=82).

Our observation window starts when the migrant arrives in the Netherlands for the first time and each individual is followed for a period of five years on a monthly time base. Thus, every trajectory is composed by a string of $(12)^5 = 60$ characters that represent the months since migration. An example is the sequence C/20-M/30-MC/10. In the example, the migrant is cohabiting with an unmarried partner for 20 months, then marries to his/her partner and finally, after 30 months marriage, becomes a parent. The number of possible combinations between these 60 months and seven states is extremely large ($7^{60}$) and thus raises problems of complexity when comparing the trajectories. Although in reality not all of these combinations are possible, the problem of complexity is still present. Therefore, we identify life course subtypes to reduce the very large number of possible combinations of the sequences into groups that can be easily interpreted.

In order to arrive at this typology of sequences, we first need to compute distances between all individual sequences. In this study, we opt for the optimal matching (OM) algorithm to do so (see Abbott & Tsay 2000). The OM algorithm measures the dissimilarity of two sequences, roughly speaking, by considering how much effort one must perform to transform one sequence into the other. In this procedure, there are three basic operations available: insertion, deletion, and substitution. Thus, when transforming one sequence into another, an item can be inserted into the sequence, an item can be deleted from the sequence or an item can be substituted by another item.
The first two operations are subsumed to so-called ‘indel’ operations. Many combinations of substitution and indel operations may be used in order to make sequences equal. The more operations necessary to make two sequences equal, the greater the dissimilarity between the sequences. Since one could argue that a substitution has a different meaning than an indel operation does, the operations can be weighted. The weights assigned to these operations are referred to as ‘substitution costs’ and ‘indel costs’. There is no optimal solution for assigning the cost of these operations and many authors agonize over this problem (see e.g. Wu 2000 for discussion). In this study we follow the commonly applied solution of using transition rates in order to get data-driven costs, thus avoiding subjectivity (Aassve et al. 2007, Piccarreta & Billari 2007). Transition rates measure the proportion of people in state $x$ at time $t$ who are in state $y$ at time $t+1$, as well as the proportion of people in state $y$ at time $t$, who are in state $x$ at time $t+1$. Indel costs are set with a reference unitary cost (either two states are the same or they are not) and substitution costs are defined as the inverse of the transition probability, assigning higher costs to less frequent transitions. All calculations were made using R’s “TraMineR” package (see Gabadinho et al. 2011).

The resulting output of the OM metric is a symmetric distance matrix that contains the minimal costs for transforming the sequences into each other. Based on this matrix, we identify more-or-less homogeneous groups of trajectories by applying the PAM (Partitioning Around Medoids) algorithm (Kaufman & Rousseeuw 1990). In this clustering method, it is necessary to specify the number $k$ of clusters in advance. Therefore, we test several cluster solutions and use the ASW (Average Silhouette Width) criterion to determine the clustering validity and to select the ‘optimal’ number of clusters (see Rousseeuw 1987). We then run a series of logistic regression models to analyze how the clusters are related to the migrant’s background characteristics. The following four predictors are included in the models: gender, age at migration, year of migration and the origin of the partner. The partner variable consists of three categories (i) Polish partner, (ii) Dutch partner, (iii) no partner. Those who spent less than three months of the sequence with a partner are classified as having no partner. When the migrants have a Dutch partner for at least three months they are classified as having a Dutch partner. Otherwise, the migrants are classified as having a Polish partner. In the case that the migrant is married but is not living at the same address as their partner, we assumed that these migrants had a Polish partner as well.
RESULTS

Distribution of states

We start with a sequential illustration of Polish young adults’ family life and migration trajectories. It is important note that the plotted state distribution (Fig. 1) does not represent individual life course trajectories, but illustrates the aggregated picture. Each individual is described for a period of five years (on a monthly time base) as being in a specific state, such as single, cohabiting, or married, and the distribution of states of the whole group is plotted for each specific month. This is therefore a descriptive approach to characterize the experiences of Polish young adults within five years after arrival in the Netherlands on a group level. The sequences are aligned on years since migration. We acknowledge that this approach obscures age-related patterns, since the start of our observation (i.e. age at migration) varies between age 20-30. Aligning the sequences on age, however, would lead to missing data since we have no information on the migrants’ family life events prior to the migration move. Therefore, we start the observation when the migrant enters the Netherlands for the first time and control for age effects in the next section.

Figure 1

Figure 1 reveals that about half of the young adult Polish labor migrants does not have a partner when they migrate to the Netherlands for the first time. The other half is more or less equally divided between those with a married and those with an unmarried partner. Note that some 25 percent of the married individuals has migrated without his/her partner, indicating that migration indeed involves spousal separation. Furthermore, it can be seen that the majority of Polish migrants does not have any children upon arrival in the Netherlands. At the end of our observation window, however, almost 60 percent of married Poles has children living in the household. Children are less common in non-marital unions. In line with previous studies (Nicolaas 2011, Dagevos 2011), we find that a large share of Polish labor migrants lives only temporarily in the Netherlands: almost 40 percent has left the country at the end of the five years of observation.

We also present the distribution plots for men and women separately in Figure 2. This figure shows that the state distribution differs strongly between men and women. Almost 60 percent of the Polish men is single when they migrate to the Netherlands, compared to about 40 percent of their female counterparts. In addition, one-third of the married men arrives in the Netherlands without their partner, while this is very uncommon among women. Women are thus more likely to
migrate and live with a partner than men. It is striking to see the difference is in particular large for non-marital cohabitation. The proportion cohabiting women is almost twice as large as the proportion cohabiting men, while the gender difference with regard to marriage is relatively small. Finally, Figure 2 shows that men are more likely to return to Poland than women are. About 40 percent of men has left the country at the end of our observation window, compared to about 30 percent among women.

Figure 2

As stated, the distribution plots illustrate aggregated family and migration states and do not give an indication about the dynamics of individual trajectories. Table 1 shows the ten most frequent trajectories for men and women separately. The representation in table 1 does not take the length of permanence in a state into account, but only the sequencing of events. These ten patterns cover about 45 percent of all sequences among men and 40 percent of all sequences among women. We first discuss the most frequent trajectories among men. As can be seen in table 1, the most frequent occurring sequence (17% of all sequences) is to remain single and return to Poland within five years. The group that remains single, but does not leave the Netherlands within our five year observation window, is the second most common. The third most frequent sequence among men are those who come with a married partner to the Netherlands and leave within five years. Thus, the three most frequent sequences among men do not include any family life transitions in the Netherlands. The fourth, fifth and sixth most frequent trajectories are those who come as single to the Netherlands and start cohabiting with an unmarried partner. It is interesting to see that none of these ten sequences includes marriage after migration to the Netherlands. Finally, two sequences in table 1 include married men who have children living in the household, while childbearing in non-marital unions is not among the ten most frequent sequences.

Among women, it can be seen that the most common trajectory (8% of all sequences) is to remain single and to leave the Netherlands within five years as well. This trajectory is, however, much less dominant than is the case among men: women who remain married and leave the Netherlands within five years are the second most common sequence covering almost 7% of all sequences. Although a substantial share of Polish women lives with a non-married partner after migration (see Fig. 2), non-marital cohabitation followed by return migration is not among the ten most frequent sequences. Married women seem thus more likely to leave the Netherlands than cohabiting women. The third most frequent ordering of events among Polish women is to live with a
non-married partner after migration and experience marriage followed by first time parenthood in the Netherlands. This pattern is more common than having children in a non-marital union, suggesting that Polish women generally prefer marriage before childbearing.

Table 1

Typology of trajectories
Since the number of possible combinations of the sequences is extremely large, we have reduced the entire set of sequences into more-or-less homogeneous subgroups by means of PAM cluster analysis. Several cluster solutions were tested (4, 5, 6, 7 cuts), of which the 6-cluster solution showed the highest average silhouette width (ASW = 0.44). We therefore specify six groups of trajectories. The state distribution within each of these six clusters is plotted in Figure 3. The figure furthermore provides information on the number and percentage of migrants who are classified within the clusters along with the ‘medoid sequence’ for each cluster. The medoid sequence is the individual sequence with the minimum average pairwise distance to all other individual sequences in the cluster (see Aassve et al. 2007). The medoid sequence is thus a real sequence, which is used as the most characteristic sequence within a particular cluster. We briefly describe each of the six clusters.

Cluster 1 accounts for 24 percent of the sample and has the medoid sequence S/60, which stands for a trajectory in which a person lives in the Netherlands without a partner during the complete observation window of five years. Some Poles in this group live together with an unmarried partner for a while, but generally short. Furthermore, some of the sequences in this cluster include return migration. These return migrants either leave the Netherlands towards the end of our observation window, or only for a short period (circular migration). The sequences in this cluster are thus all characterized by living single in the Netherlands for a relative long time.

Cluster 2, which compromises the largest group of the sample (27%) differs from the first in that these migrants all leave the Netherlands early within our five year observation window. This is reflected in the medoid sequence of S/20-R/40. A person who experienced such a trajectory lived for 20 months without a partner in the Netherlands and has then left the country and did not return to the Netherlands for the remaining 40 months of observation. The majority of the migrants in this cluster does not have a partner during the time of residence in the Netherlands (see Figure 3). Those with a partner are about equally divided between married and cohabiting individuals. Around half of the married individuals in this cluster lives without their partner in the Netherlands and returns to
their partner in Poland. It is interesting to see that hardly any of the migrants in this cluster have children living in the household.

The sequences in cluster 3, which contain 16 percent of the sample, are characterized by the medoid S/14-C/46, implying that the migrant is single upon arrival in the Netherlands and starts cohabiting with an unmarried partner after 14 months until the end of our observation window. Although some individuals in cluster 3 become married and/or have children living in the household towards the end of the five years of observation (see Fig. 3), all of these trajectories include a long period of non-marital cohabitation. Cluster 4 is also characterized by a long period of non-marital cohabitation, but then with children living in the household. This group is characterized by the medoid S/6-C/2-CC/52, which stands for a trajectory of living single for six months in the Netherlands, then start cohabiting for 2 months, followed by cohabiting and a child living in the household. This cluster is the smallest cluster of the sample (7%), suggesting that children in non-marital unions are less common.

In cluster 5, the sequences are characterized by a relative long period of living with a married partner in the Netherlands. The medoid sequence of this group is MP/51-R/9, which means that the migrant lives with a married partner for 51 months and then leaves the Netherlands until the end of our observation. This cluster (15%) is slightly smaller than cluster 3 (16%), which was characterized by a long period of non-marital cohabitation (see also Figure 1). There are two interesting differences between cluster 3 (cohabiting group) and cluster 5 (married group). First, the vast majority of the married Poles migrates with their partner to the Netherlands, while the cohabiting Poles often arrive as single in the Netherlands. Second, a large share of the married cluster leaves the Netherlands within five years, whilst return migration is uncommon in the cohabiting cluster.

The last cluster contains trajectories that are characterized by the medoid sequence MP/5-MC/55, which means that the migrant has lived with a married partner for five months and then becomes a parent and remains in the Netherlands for another 55 months. This cluster thus represents sequences that are marked by a long period of living with a married partner and children in the household. The cluster compromises around 11 percent of the sample and is therewith larger than cluster 4 (7%), which includes long periods of non-married cohabitation and children living in the household. A similarity between cluster 6 and cluster 4 is that return migration is uncommon in both of these clusters, suggesting that the presence of children in the household decreases the propensity to leave the Netherlands.

*Figure 3*
Now that we have described each of the clusters, we test how these clusters differ in compositional characteristics by means of a series of logistic regression analyses (see Table 2). Note that there are no odds ratios for having no partner in clusters 3, 4, 5 and 6, since all trajectories in these clusters include at least three months of living with a partner. We start with the results for cluster 1, which is characterized by a long period of living single in the Netherlands. The results confirm the findings in Figure 2, that Polish men are more likely to live without a partner in the Netherlands than their female counterparts. Furthermore, having no partner during the time of residence in the Netherlands is obviously positively related to the likelihood of membership in cluster 1. We also find that those with a Dutch partner are significantly more likely to be grouped in this cluster than those with a Polish partner, which is probably due to the fact that this cluster also contains short periods of non-marital cohabitation (see Fig. 3). As can be seen in Table 2, we do not find an association between the age of observation and the likelihood of living single in the Netherlands, as none of the age-effects are statistically significant.

With regard to the results for cluster 2 (Table 2), which was characterized by an early departure from the Netherlands, we find that men are significantly more likely to follow this type of trajectory than women. Confirming common sense, Polish migrants with a Dutch partner have a lower likelihood to leave the Netherlands than those with a Polish partner. Having no partner during the time of residence in the Netherlands strongly increases the likelihood of return migration. The effect of the age of observation shows an U-shaped relationship. The migrants who were observed between age 25-30, 26-31, and 27-32 are significantly less likely to leave the Netherlands than the reference category (age 20-25). The differences between the older age groups and the reference category are not significant.

Clusters 3 and 4 can be summarized as cohabiting trajectories, in which one form is without (cluster 3) and the other is with children living in the household (cluster 4). As can be seen in Table 2, Polish men are less likely to cohabit with an unmarried partner than women. This gender difference is even stronger with regard to cohabitation with children in the household. Furthermore, Table 2 shows that Poles with a Dutch partner are more likely to cohabit than those with a Polish partner. Cohabiting without children (cluster 3) is more likely among the younger migrant groups, while cohabitation with children is more likely among the oldest migrant group (30-35). Note that the fact that we do not find a strong association between age of observation and the trajectories of cluster 4 may be due to the lower sample size of this cluster (N=374).
The last two clusters (5 and 6) differ from the previous two in that these clusters include trajectories with long periods of living with a married partner instead of an unmarried partner. Table 2 shows that there is no difference between men and women in the likelihood of living with a married partner. Contrary to the cohabiting clusters, Polish migrants with a Dutch partner are less likely to be married than those with a Polish partner. Finally, there is no significant effect of age of observation with regard to the trajectories in cluster 5. Older migrants are however much more likely to live with a married partner and children in the household (cluster 6).

Table 2

CONCLUSION AND DISCUSSION
To be completed.

FUNDING
This work is part of and financed by the European Research Council, Starting Grant project (number 263829) “Families of Migrant origin: a Life course perspective” (FaMiLife) awarded to H.A.G. de Valk.
REFERENCES


**Figure 1.** Transversal distribution of family life states and return migration

**Figure 2.** Transversal distribution of family life states and return migration, by gender
Figure 3. Transversal distribution of family life states and return migration, 6 life course clusters
Table 1. Ten most frequent trajectories with the same sequencing of events

<table>
<thead>
<tr>
<th></th>
<th>Sequence</th>
<th>Percent</th>
<th></th>
<th>Sequence</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>S – R</td>
<td>17.1</td>
<td></td>
<td>S – R</td>
<td>7.9</td>
</tr>
<tr>
<td>2</td>
<td>S</td>
<td>6.3</td>
<td></td>
<td>MP – R</td>
<td>6.5</td>
</tr>
<tr>
<td>3</td>
<td>MP – R</td>
<td>4.6</td>
<td></td>
<td>C – MP – MC</td>
<td>4.1</td>
</tr>
<tr>
<td>4</td>
<td>S – C – S</td>
<td>3.3</td>
<td></td>
<td>S</td>
<td>3.8</td>
</tr>
<tr>
<td>5</td>
<td>S – C – S – R</td>
<td>3.1</td>
<td></td>
<td>S – C</td>
<td>3.7</td>
</tr>
<tr>
<td>6</td>
<td>S – C</td>
<td>2.8</td>
<td></td>
<td>C – MP</td>
<td>3.5</td>
</tr>
<tr>
<td>7</td>
<td>MP</td>
<td>2.5</td>
<td></td>
<td>C</td>
<td>3.0</td>
</tr>
<tr>
<td>8</td>
<td>MP – MC</td>
<td>1.9</td>
<td></td>
<td>MP</td>
<td>3.0</td>
</tr>
<tr>
<td>9</td>
<td>C – S – R</td>
<td>1.8</td>
<td></td>
<td>C – CC</td>
<td>2.6</td>
</tr>
<tr>
<td>10</td>
<td>MC</td>
<td>1.8</td>
<td></td>
<td>MP – MC</td>
<td>2.5</td>
</tr>
</tbody>
</table>

Note: S = single; C = cohabiting; CC = cohabiting + child; MP = married partner; MC = married + child; R = return

Source: Social Statistical Database (SSD)
<table>
<thead>
<tr>
<th></th>
<th>Cluster 1</th>
<th>Cluster 2</th>
<th>Cluster 3</th>
<th>Cluster 4</th>
<th>Cluster 5</th>
<th>Cluster 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.17***</td>
<td>0.29***</td>
<td>0.35***</td>
<td>0.13***</td>
<td>0.22***</td>
<td>0.07***</td>
</tr>
<tr>
<td>Male</td>
<td>1.60***</td>
<td>1.25**</td>
<td>0.78**</td>
<td>0.41***</td>
<td>0.85</td>
<td>0.99</td>
</tr>
<tr>
<td>Partner</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Polish (ref)</td>
<td>1.45*</td>
<td>0.68**</td>
<td>2.06***</td>
<td>1.58**</td>
<td>0.24***</td>
<td>0.62**</td>
</tr>
<tr>
<td>Dutch</td>
<td>2.91***</td>
<td>4.78***</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Single a</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Age of observation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-25 (ref)</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>21-26</td>
<td>1.30</td>
<td>0.88</td>
<td>0.82</td>
<td>0.87</td>
<td>1.17</td>
<td>0.96</td>
</tr>
<tr>
<td>22-27</td>
<td>1.41</td>
<td>1.02</td>
<td>0.79</td>
<td>0.74</td>
<td>0.92</td>
<td>1.09</td>
</tr>
<tr>
<td>23-28</td>
<td>1.26</td>
<td>0.83</td>
<td>0.83</td>
<td>1.03</td>
<td>0.96</td>
<td>1.72</td>
</tr>
<tr>
<td>24-29</td>
<td>1.43</td>
<td>0.81</td>
<td>0.84</td>
<td>0.64</td>
<td>1.04</td>
<td>1.63</td>
</tr>
<tr>
<td>25-30</td>
<td>1.36</td>
<td>0.65*</td>
<td>0.72</td>
<td>1.06</td>
<td>1.20</td>
<td>2.11*</td>
</tr>
<tr>
<td>26-31</td>
<td>1.19</td>
<td>0.62**</td>
<td>0.83</td>
<td>1.18</td>
<td>1.06</td>
<td>2.65**</td>
</tr>
<tr>
<td>27-32</td>
<td>1.24</td>
<td>0.59**</td>
<td>0.78</td>
<td>1.10</td>
<td>1.15</td>
<td>2.71**</td>
</tr>
<tr>
<td>28-33</td>
<td>1.07</td>
<td>0.75</td>
<td>0.45***</td>
<td>1.53</td>
<td>0.82</td>
<td>4.46***</td>
</tr>
<tr>
<td>29-34</td>
<td>0.81</td>
<td>0.71</td>
<td>0.65*</td>
<td>1.39</td>
<td>0.82</td>
<td>4.13***</td>
</tr>
<tr>
<td>30-35</td>
<td>0.99</td>
<td>0.90</td>
<td>0.53**</td>
<td>1.80*</td>
<td>0.79</td>
<td>3.21***</td>
</tr>
<tr>
<td>Year of migration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004 (ref)</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>2005</td>
<td>0.79*</td>
<td>1.21*</td>
<td>0.89</td>
<td>0.78</td>
<td>1.42**</td>
<td>0.92</td>
</tr>
<tr>
<td>2006</td>
<td>0.85</td>
<td>0.79*</td>
<td>1.09</td>
<td>1.05</td>
<td>1.46***</td>
<td>1.08</td>
</tr>
<tr>
<td>2007</td>
<td>0.88</td>
<td>1.31*</td>
<td>1.10</td>
<td>0.80</td>
<td>0.68*</td>
<td>1.04</td>
</tr>
<tr>
<td>No. of observations</td>
<td>1273</td>
<td>1458</td>
<td>876</td>
<td>374</td>
<td>796</td>
<td>578</td>
</tr>
<tr>
<td>Nagelkerke $R^2$</td>
<td>0.08</td>
<td>0.15</td>
<td>0.16</td>
<td>0.07</td>
<td>0.16</td>
<td>0.14</td>
</tr>
</tbody>
</table>

Source: Social Statistical Database (SSD)

* $p < .05$; ** $p < .01$; *** $p < .001$

a Variable omitted in models 3, 4, 5, 6 since all sequences in these clusters contain periods of living with a partner