OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 0305–9049
doi: 10.1111/obes.12295

Immigrants' Wage Growth and Selective Out-Migration*

Govert E. Bijwaard†,‡ and Jackline Wahba‡,§

†*Netherlands Interdisciplinary Demographic Institute, NIDI-KNAW/University of Groningen, 2502 AR,The Hague, The Netherlands (e-mail: bijwaard@nidi.nl)*
‡*IZA, Bonn, Germany*
§*University of Southampton, Southampton, SO17 1BJ, UK (e-mail: j.wahba@soton.ac.uk)*

Abstract

This paper examines immigrant wage growth taking selective out-migration into account using administrative data from the Netherlands. Addressing a limitation in the previous literature, we address the potential endogeneity of immigrants' labour supply and out-migration decisions on their earning profiles using a correlated competing risk model. We distinguish between labour and family migrants, given their different labour market and out-migration behaviours. Our findings show that accounting for selective labour supply is as important as accounting for selective out-migration. Controlling only for out-migration selectivity would underestimate immigrants' wage growth, whilst controlling only for labour market selectivity would overestimate their wage growth. This shows that different selections are important for different types of migrants.

I. Introduction

How do immigrants fare in the host country? Both academics and policy makers are interested in immigrants' economic performance in the host country, and immigrants' earnings are commonly used as a measure of their productivity and contribution to the economy. It is therefore not surprising that some host countries use earnings as a measure to identify 'the best and the brightest'. For example, both the Netherlands and the UK use earning thresholds to identify highly skilled immigrants, and more recently, the UK introduced further earning threshold requirements for immigrants wishing to settle permanently in the UK.

Following Chiswick (1978) seminal work, a sizable body of literature has examined immigrant performance in the host country, focusing on immigrant earnings and their growth. This literature has sufficiently evolved to deal with several empirical challenges,

JEL Classification numbers: F22, J61, C41.
*Financial support from the NORFACE research programme on Migration in Europe – Social, Economic, Cultural and Policy Dynamics and the Economic and Social Research Council (RES-167-25-0678) is gratefully acknowledged. We thank Statistics Netherlands, and Han Nicolaas in particular, for data support. We thank the participants of the 7th Annual Conference on Immigration in OECD Countries for their helpful comments.

such as arrival cohort effects (see Borjas, 1985) and time effects (e.g. Anteco, Kuhn and Trejo, 2006). Another challenge has been dealing with the non-randomness of out-migration given the limitation of available data. Although recent studies using longitudinal data are able to control for unobserved heterogeneity, such data still pose a limitation in terms of identifying the wage growth of immigrants who remained in the host country. Not only is return migration inferred rather than observed and is typically confounded with attrition given the absence of administrative data collected by immigration authorities, but selection is also assumed to occur on time-constant unobservables (see Dustmann and Görlach, 2015 for an excellent review of this literature).

Although this literature has developed to examine immigrants' performance in the US (see e.g. Hu, 2000; Lubotsky, 2007; Biavaschi, 2016) and in many other countries where panel data and longitudinal data exist (e.g. Pischke, 1992), existing studies have not addressed the potential endogeneity of wages, labour supply and out-migration. Indeed, immigrants may leave because of their wage levels or wage growth. While the unsuccessful ones could be those who leave due to a lack of employment and negative wage growth, it could also be the successful migrants who decide to leave, having experienced positive wage growth and reached their target savings. We address this challenge which has hitherto been overlooked in this literature – namely, exploring the notion that out-migration might be correlated with the level of immigrant earnings. Hence, our focus in this paper is on how failing to account for out-migration selectivity and wage endogeneity would bias the estimates of immigrant earnings growth.[1]

Our main contribution is to address this potential endogeneity between wages/employment and the out-migration choice when estimating immigrant earnings. We take into account the previously unexplored interdependence between three important factors, namely, the labour market status of the migrant since arrival, his/her wages in the country of immigration and the out-migration decision, as they can all affect the estimated wage growth of current immigrants. Unlike existing literature and noted as a strong assumption by Dustmann and Görlach (2015) in their recent review, we do not assume that immigrants' decisions on labour supply are not determined by their out-migration plans. In particular, we study what the growth in mean earnings of immigrants who arrived between 1999 and 2007 and of those who stayed in the host country would have been if there had been no selective out-migration. We use administrative data from the Netherlands to understand how selective out-migration as well as the endogeneity of the labour market status and wages influence immigrant wage growth. We observe all immigrants who entered the country between 1999 and 2007, as well as their motive for migration: whether for labour migration or otherwise, the timing of return in case of out-migration, the exact detailed (monthly) information on their labour market status and their income since their arrival.

Furthermore, different types of migrants, for example, labour or family migrants, behave differently in the labour market and have different migration behaviour (see Bijwaard, 2010). Census data and standard panel surveys do not distinguish between migrant types. Hence, previous studies have been unable to distinguish between the behaviour of different types of migrants, rather examining them as a homogeneous group. In the Netherlands – as in other European countries – family migrants represent a higher proportion of total

[1] We do not compare immigrant and native earnings.

migrants compared with labour migrants. For example, about 31% of all non-Dutch immigrants in the age group 18–64 are labour migrants, while 34% are family migrants. Our data allow us to distinguish between these two potentially different types of migrants who might experience different wage trajectories and different patterns of out-migration selection, and whose aggregation into one group might therefore potentially create biased results.

Our empirical strategy is to adopt a correlated multi-state model (full-CCRM) to estimate immigrant earnings growth and address the endogeneity of wages, labour market status and out-migration selection. However, we estimate several models to disentangle the importance of out-migration selectivity from the endogeneity of the labour market. Hence, we first assume that we only observe retrospective data on the migrants still in the host country and we start by estimating naive cross sectional models, and we then apply 'standard' panel data methods using either random or fixed effects, as previously undertaken in the literature by using similar (constructed) data. This allows us to benchmark our results. We then estimate a correlated multi-state model (labour-CCRM) that takes the endogeneity of wages and labour market status into account, but not the out-migration. We subsequently use the information on the timing of out-migration and estimate a model with selection into out-migration, albeit ignoring the labour market transitions. Our final model combines these last two models into a multi-state model (full-CCRM) that takes the out-migration selectivity as well as the endogeneity of the labour market status into account. We also check the robustness of our results by estimating the earnings profiles of the 1999 arrival cohort to control for any cohort effects. Our correlated multi-state model is closely related to the multivariate Mixed Proportional Hazards model with endogenous wage formation for job seekers and their subsequent wage of Gaure, Røed and Westlie (2012). They use this modelling strategy to identify the causal effects of unemployment duration and active labour market policy participation. Their model, and ours, builds on the timing-of-events approach of Abbring and van den Berg (2003).

Our findings show that while classic panel data models overestimate immigrant earnings growth, controlling only for labour market selectivity would also overestimate the wage profiles of migrants as the selectivity of out-migration is ignored. On the other hand, controlling only for out-migration selectivity would underestimate the wage profile, since it ignores the labour market transitions and periods of non-employment. Indeed, our results highlight the importance of taking into account both the endogeneity of the labour supply and out-migration when estimating immigrant earnings profiles. Accounting for selective labour market transitions seems as important as accounting for selective out-migration.

The results also show that labour migrants behave differently from family migrants. For labour migrants, controlling for migration selectivity is more important earlier on in the migration cycle whilst controlling for labour market endogeneity tends to be more important in later years. For family migrants, both selectivities are important, as ignoring selective out-migration would overestimate the wage growth whilst ignoring the endogeneity of labour supply would underestimate the wage growth. Our results further suggest that migration motives are more important for the wage growth than gender differences. Treating the migrants as one group would hide these differences.

The remainder of this paper is structured as follows. In section II, we summarise the main theoretical and empirical literature on immigrant earnings growth, before presenting the data and discussing the main characteristics of recent migration to the Netherlands in

section III. Section IV discusses the different estimation methods that we apply to estimate the immigrant earnings growth. Section V provides the estimated impact of years since migration on the immigrant earnings profile for all these models. Section VI distinguishes between various migrant groups, presenting the estimates for labour vs. family migrants as well as comparing the findings by main region and country of origin, and by gender. We also check the robustness of our results using one cohort only, the wage growth of migrants who arrived in 1999, to control for any potential cohort effect. Finally, the last section summarises the findings.

II. Theoretical and empirical background

The success of immigrants in the labour market is commonly measured by their earnings. A key issue in the debate on migration is whether immigrants contribute to the economy. In this respect, the evolution of immigrants' income over their migration cycle is an important indicator of their productivity and economic performance.

When they first arrive in the new host country, immigrants often lack country-specific human capital- such as knowledge about the labour market and language skills. More time spent in the host country is assumed to lead to more human capital investment, which is likely to increase the immigrants' productivity, enabling them to progress up the earnings ladder. The effect of years since migration is expected to have a positive effect on immigrant earnings, given such investments in host country-specific human capital. This investment process explains the relatively rapid growth rates of immigrant earnings observed in cross-sectional studies. For instance, Chiswick (1978) provides the first insights into this relationship using cross-sectional estimates, albeit assuming that the unobserved characteristics of the immigrants are stable over time. However, the positive impact of the years-since-migration variable in cross-sectional earnings equations captures both the quality of earlier immigrant cohorts as well as the increase in the host-specific human capital. Subsequent studies for example, by Borjas (1985) and LaLonde and Topel (1992), used repeated cross sections to identify the cohort effect, albeit assuming that the time effects are the same for natives and immigrants. Nonetheless, repeated census data remain problematic as the estimate of the cohort fixed effects is a mixture of earnings of those who stay and those who leave.

Several studies have resorted to using a longitudinal stock-based sample (see e.g. Hu, 2000; Lubotsky, 2007) and have typically found smaller immigrants' wage growth compared to estimates from repeated cross-sectional data (see e.g. Borjas, 1985). Recent papers by Abramitzky, Boustan and Eriksson (2014) and Biavaschi (2016), try to circumvent the out-migration selectivity challenge by using repeated census data of both sending and host countries.[2] However, this is problematic if not all migrants (returnees) are identified in the host (sending) country census and different definitions of migrant/returnee or education are used in the sending and host country. Other studies have used panel data, such as Pischke (1992), or administrative data, such as Edin, Lalonde and Åslund (2000) and Barth, Bratsberg and Raaum (2004). However, the major challenge that all such studies

[2] See Lubotsky (2007) for a discussion on the limitations of longitudinal data in estimating immigrants' earnings trajectories.

face in tackling out-migration selectivity is the absence of administrative data collected by immigration authorities, where information on out-migration is observed rather than inferred from attrition seen in survey samples. This is certainly the case in the Netherlands, where all immigrants have to register. This paper uses administrative data where all immigrants and out-migrants are observed and hence it is able to control for the selectivity of out-migration when estimating the average earning profiles of immigrants in the host country, as well as what would have been the average earning profiles of immigrants if there were no out-migration selection.

The level of out-migration has been high for both the US and many European countries. Jasso and Rosenzweig (1982) report that between 20% and 50% of legal immigrants re-emigrated from the US in the 1970s. Bijwaard (2010) reports return probabilities between 25% and 50% within five years since immigration – depending on the migration motive – for migrants entering the Netherlands during the period 1995–2003. Similar numbers have been found for other Western European countries (see e.g. Edin *et al.*, 2000 for Sweden; Jensen and Pedersen, 2007 for Denmark; and Dustmann, 1995 for Germany). When many migration decisions are temporary and migrants who re-migrate are more (less) successful than those remaining in the country, ignoring such selective out-migration will under- or over-state the economic performance of the original immigrants who have stayed. Dustmann and Görlach (2015) provide an excellent review of the potential biases, identifying assumptions associated with the various types of data adopted in the literature, as well as their impacts on the estimates of immigrant earning profiles.

Temporary migration can be planned and part of an optimal strategy to maximise lifetime utility. Out-migration could be due to higher preference for consumption in the own country relative to that in the host country e.g. Galor and Stark (1991). Also, as Dustmann and Weiss (2007) show, return migration may be motivated by lifetime utility that includes consumption and locationally fixed factors that are complementary to consumption or differences in relative prices in host and home country. Temporary migration allows the migrant to take advantage of high wages abroad and low prices at home. Individuals migrate for a period of time where wages are higher so that they can accumulate savings overseas that they can use to set up businesses upon return or acquire skills that are highly rewarded in the source country. Within this framework, better conditions at home would also reduce the benefits for migrants from staying overseas. Another mechanism behind temporary migration might be due to failure, when the migrant realises that their expected value of lifetime income is lower than that in the origin country and they return. In this case the decision is unplanned and arises as a result of either imperfect information about the host country in terms of labour market prospects or the cost of living, or the inability to fulfil the migration plans in terms of target savings. This kind of out-migration is expected to take place early on in the migration cycle, see Borjas and Bratsberg (1996). In essence temporary migration can be explained in a framework based on the selection model of Roy (1951), in which the composition of migratory flows depends on the relative distribution of incomes between the home and host countries, and average returns on human capital. Hence, out-migration can be both planned, therefore leading to positive selection of out-migrants, and unplanned, perpetuating negative selection.[3] In this paper, we argue that

[3] See Wahba (2014) for a review of this theoretical and empirical literature on return migration.

an additional complication is that realised immigrant earnings are endogenous with out-migration (because decisions related to the labour market are usually made in conjunction with out-migration decisions), and thus make the estimation of immigrant wage profiles and out-migration selection more complicated.

This is an important empirical challenge that has not been addressed in the immigrant wage growth literature to date. As noted by Dustmann and Görlach (2015, p. 502), a strong assumption made in almost all the literature is that immigrants' decisions on labour supply and job choices are not determined by their out-migration plans. To our knowledge, the only paper that studies the influence of individual labour market changes on the return decision is Bijwaard and Wahba (2014), which adopts a different focus, namely whether high- or low-income immigrants leave sooner. A related paper by Bijwaard, Schluter and Wahba (2014) uses a timing-of-events model to control for the endogeneity of labour market changes in examining the impact of unemployment on return migration. They find that unemployment induces labour migrants to return, while re-employment renders the migrants more prone to staying. However, neither of these two papers study wage growth. Indeed, our current paper's contribution lies in adopting a correlated multistate model (CCRM) to estimate immigrant earnings growth and address the endogeneity of wages, labour market status and out-migration selection, which has not been previously addressed.

III. Administrative panel data on the population of immigrants to the Netherlands

This paper makes use of the unique feature of the administrative data in the Netherlands where all legal immigration by non-Dutch citizens to the Netherlands is registered in the Central Register of Foreigners (Centraal Register Vreemdelingen, CRV), using information from the Immigration Police (Vreemdelingen Politie) and the Immigration and Naturalisation Service (Immigratie en Naturalisatie Dienst, IND). It is mandatory for every immigrant to notify the local population registrar immediately upon arrival in the Netherlands if he or she intends to stay for at least two-thirds of the forthcoming six months.[4] The data comprise the entire *population* of immigrants who entered during our observation window of 1999–2007, whereby after merging it with other administrative registers we obtain a panel.

In addition to the date of entry and exit, the CRV also records the migration motive of the individual. The motive is coded according to either the immigrant's visa status, or the immigrant reports the motive upon registration in the population register. Statistics Netherlands distinguishes between the following migration motives: labour, family, study, asylum seekers (and refugees), and others. EU-citizens are required to register in the Netherlands, just as natives are. See Bijwaard (2010) for an extensive descriptive analysis of the various migration motives. We focus on the two main migration motives, namely

[4]Although some migrants may not officially inform the authorities that they are about to leave the Netherlands, all citizens (immigrants and natives) are required to register with their municipalities (this is a prerequisite for many social services, and for tax-benefit matters). Thus, the Dutch authorities treat any migrant who has no entries in the tax-benefit register and does not appear in the register of another municipality to have left the country. Hence, migrants who fail to de-register are periodically identified and removed from the registers by the authorities in a step labelled 'administrative removal'. We know the date the migrant was last observed on the register and the date they were administratively removed.

labour and family migrants. In particular, about 31% of all non-Dutch immigrants aged 18 and older who arrived in 1999–2007 are labour migrants, while 34% are family migrants.[5] As it is possible that the official migration motive does not always match the migrant's true intention. This has induced Statistics Netherlands to define labour migrants as only those with a labour motive who are employed in the Netherlands within three months of their entry.[6] We also limit our analysis to the first migration spell, i.e. we do not include repeat migrants who left the Netherlands and subsequently returned.

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

Our data benefits from enabling us to observe immigrants from entry until they leave or until the end of the observation period (1/1/2008). Although our migration duration has a maximum of nine years, our results are still very informative since we compare the estimates of various models over the same time period and before the onset of the global economic crisis in 2008. It could also be argued that most out-migration occurs within the first few years after arrival and for most immigrant cohorts the early years are observed in our data.

Historical and institutional background

In the early 1960s, the Netherlands changed from an emigrant to an immigrant country, whereby its immigration has followed the general European immigration patterns of post-World War II and postcolonial immigration: unskilled manpower recruitment and the arrival of refugees. The first period is characterised by the de-colonisation of Indonesia in 1949, as a consequence of which 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 via bilateral agreements with the main countries. The total inflow of immigrants reached 235,000 in the 1970s. 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. Starting in the 1980s, immigration

[5] About 16% of these migrants are students, 13% are refugees and 6% are migrants for other reasons. Most of the international student migrants to the Netherlands return home, usually right after graduation, and they do not have access to Dutch labour market, see Bijwaard and Wang (2016). Refugees are only registered and allowed to work after receiving a permanent residence permit. This may take up to two years. This implies we do not know their exact length of stay or their full labour market behaviour.

[6] A point of concern might be that this selection rule imposed by Statistics Netherlands to identify labour migrants affects the estimated wage profiles. Bijwaard *et al.* (2014) have tested the impacts of varying cut-off points of this selection rule of labour migrants on the impact of unemployment and re-employment on the out-migration hazard and only found minor changes. We are therefore confident that the selection rule does not influence the estimated wage profiles.

is characterised by the family reunification/formation of 'guestworkers'. Additionally, the flow of political refugees and asylum seekers has increased dramatically. In the political discourse it is often forgotten that the number of labour immigrants from neighbouring countries and other EU countries has always been substantial. In the last twenty years, the majority of labour immigrants came from these countries or from other western countries. The forming of the European Union and the EU treaty of 1993 that allowed free movement of people within the union has facilitated the migration within the EU. In 2004 the EU expanded with the addition of 10 more countries.[7] However, only in May 2007 did the citizens coming from these new EU countries receive full access to the Dutch labour market. In 2004, in light of the Lisbon agreement, the Dutch government recognised the need for highly skilled migrants to sustain further economic growth, a so-called 'knowledge regulation' that simplifies entry into the Netherlands for migrants who will earn more than €47 thousand a year.[8]

The annual emigration from the Netherlands was rather stable from the late 1950s until the late 1980s. In that period, around 60 thousand people left the country each year. In the early 1950s emigration had peaked at 80 thousand people, which was mainly due to active emigration policies of the Dutch government. These active emigration policies were reverted to an active immigration policy in the 1960s when a shortage of labour occurred. In the 1990s and early in this century, emigration increased rapidly to reach a new peak at 132 thousand emigrants per year. However, the composition of the recent emigrant cohort differs substantially from the composition of that of the 1950s and 1960s. In the latter period the emigrants were almost entirely native Dutch, while two-thirds of the recent emigrants are non-native. Recent research has shown that many migrants leave fast, within five years 40% of the recent migrants having left the country, and that migration experience accelerates this process, see Bijwaard (2010), Bijwaard and Wahba (2014) and Bijwaard *et al.* (2014).

It is important to briefly describe the institutional background regarding immigrants' access to the Dutch labour market. Any immigrant from the EU15 can move freely in the Dutch labour market, as can, since 2004, immigrants from the new EU and countries that joined the EU after 2004, except for Bulgarians and Rumanians. All non-EU migrants need a work permit. Migrants from less-developed countries (LDCs) and developed countries (DCs) differ in that immigrants from DCs (EEA, Australia, Canada, Japan, New Zealand, USA, South-Korea and Switzerland) are exempted from obtaining a work permit before entry.[9]

Immigrants from the EU15 rapidly gain the same employment and benefits' rights as natives, essentially after three months of full-time employment. By contrast, all other immigrants have to gain permanent residency before gaining these rights, the qualifying period for which is five years of continuous full-time employment. Having lost a job, non-EU immigrants without a permanent residence permit have three months to find another job before they lose the right to stay in the Netherlands.[10]

[7] The enlargement in 2007 with five more countries is beyond the observation period of our database.

[8] See Zorlu and Hartog (2002) and Van Ours and Veenman (2005) for a more detailed discussion on the immigration to the Netherlands.

[9] See online Appendix A for more details.

[10] Note that Bijwaard *et al.* (2014) find that unemployment triggers out-migration, but there is no significant increase in out-migration around the five-year period.

A foreigner who wants to stay in the Netherlands for the purpose of family reunification or family formation has to apply for a temporary residence permit. For family migrants the same rules apply as for labour migrants. The minimum age requirement for family reunification in the Netherlands is 18 years, and for family formation, 21 years. The sponsor can be someone with Dutch nationality, or with a foreign nationality with a residence permit, but the sponsor must prove to have sufficient income that is available for at least one year. The required level of the income for family reunification is 100% of the level of social assistance for a couple and for family formation it is the minimum wage (which was raised in 2004 to 120% of social assistance).

A residence permit for family reasons is in principle always granted for a temporary period of one year first. This temporary residence permit can be renewed if the conditions are still being fulfilled. After three years the dependent spouse or partner of a Dutch national or a foreign national with a permanent residence status can apply for an independent temporary resident permit. After five years of legal stay someone can apply for a permanent residence status. Passing an integration exam (in case someone did not pass this exam for the independent residence status) and sufficient resources are conditions for this residence permit for indefinite duration.

All migrants are obligated to register in the municipal register of population (GBA) with the municipality where they have taken up residence. The GBA is a database that keeps the personal data of all inhabitants in the Netherlands. Government services who need personal data for their duties receive this information from the GBA. Registration in the GBA is mandatory for anybody working in the Netherlands, as many tax registers use this information. Failure to register can have severe legal consequences.

Descriptive statistics

Our data comprise 194,775 immigrants, including 92,893 labour immigrants and 101,882 family immigrants. Almost 53% of our sample are women, 64% are married and 70% are younger than 35 years of age. Upon entry, 51% were employed (all labour migrants and 5% of the family migrants). At the last point when a migrant is observed – either at the end of our observational period on 1/1/2008 or at the time of emigration – about 50% are employed (67% of the labour migrants and 35% of the family migrants). At the time of emigration, only 37% are employed (labour 49% and family 12%).

We use real gross wages deflated by the official Dutch price deflator throughout the analysis. Table 1 shows immigrant earnings' growth over time spent in the host country for all immigrants and distinguishes between immigrants who stayed (stayers, i.e. those who remained until 1/1/2008) and those who out-migrated (emigrants, i.e. those who left the Netherlands before 1/1/2008). We report the average income for all migrants in columns (1), (4) and (7), the average income for the migrants with positive wage earnings, excluding non-wage earners migrants with zero income –in columns (2), (5) and (8) and, the average income for the migrants who are continuously employed –in columns (3), (6) and (9). The first row shows the average income at the time of arrival in the Netherlands. Since not all immigrants are employed upon arrival (when first observed for the first time in the data), we report in the third row the average income three months after arrival. We use the wage three months after arrival as the initial reference wage in all our analysis below, and we

Table 1

Evolution of average real income

	All migrants			Stayers‡			Emigrants§		
	All	Emp*	Emp† Cont	All	Emp*	Emp† Cont	All	Emp*	Emp† Cont
Entry	€1,240	€2,951	€2,939	€910	€2,314	€2,667	€1,895	€2,776	€3,401
No. of migrants	192,986	86,650	51,140	128,191	47,019	32,184	64,795	39,631	18,956
After 3 months	€1,510	€2,946	€3,539	€1,149	€2,609	€3,268	€2,215	€3,413	€3,981
Last observation	€1,580	€3,077	€4,226	€1,554	€2,633	€3,798	€1,631	€4,391	€4,953
Observation end	€1,554	€2,634	€3,798	€1,554	€2,633	€3,798	–	–	–
Years since migration									
0–1 year	€1,524	€2,527	€3,483	€1,197	€2,156	€3,158	€2,171	€3,132	€4,036
1–2 year	€1,572	€2,743	€4,089	€1,264	€2,191	€3,536	€2,203	€3,904	€4,934
2–3 year	€1,504	€2,642	€4,332	€1,269	€2,136	€3,683	€2,115	€4,147	€5,496
3–4 year	€1,441	€2,505	€4,384	€1,296	€2,144	€3,835	€1,965	€4,132	€5,854
4–5 year	€1,376	€2,376	€4,174	€1,297	€2,135	€3,773	€1,777	€4,052	€5,809
5–6 year	€1,378	€2,333	€4,109	€1,327	€2,142	€3,718	€1,751	€4,447	€6,733
6–7 year	€1,417	€2,333	€4,151	€1,404	€2,231	€3,939	€1,569	€4,227	€6,671
7–8 year	€1,429	€2,320	€4,002	€1,430	€2,263	€3,868	€1,406	€4,265	€7,044
8–9 year	€1,437	€2,341	€4,298	€1,433	€2,304	€4,158	€1,633	€5,703	–

Notes: *Only for wage-earners employed migrants, employed at observation time.† Only for wage-earners continuously employed migrants.‡ Migrants who remained in the Netherlands until the last observation time, 1/1/2008.§ Migrants who left the Netherlands before the last observation time.

check the robustness of our results to the use of that initial wage in section 'findings by gender and migration motive'. The fourth row presents the last observed income, which includes the last income of those who left. Row five displays the income at the end of our observation window on 1/1/2008, and thus captures stayers only. The next nine rows show the average income for years since migration (YSM).

There are a number of interesting observations to note. Looking at all immigrants suggests that income at entry is lower than the last observed income and income before out-migration. When restricting the sample to wage earners as our group of interest, the average wage for those migrants at the time of last observation is higher than for those still in the country at end of the observation date (column 2). Exploring this further indicates that the average income of stayers is lower than that of emigrants, and those who are continuously employed for both groups earn more. However, interestingly, when considering all stayers and all emigrants – i.e. not restricting to positive wage earners – there seems to be a large difference. Hence, as we hypothesise, although out-migration selection is important, the role played by the labour market is also likely to affect the out-migration selection. Indeed, we investigate this issue in a more rigorous manner throughout the rest of this paper. It is important to note that these descriptive statistics confound both the cohort and time effects, as well as the selectivity of out-migration and the potential endogeneity of wages and out-migration, which we will deal with in the next section.

In order to understand better the role of the labour market, Table 2 depicts the employment dynamics for the migrants. About a quarter of the migrants are never employed (all family migrants). Although the emigrants more often have employment experience, they less often have more than one employment spell. About 80% of migrants never receive

Table 2

Employment dynamics: number of spells in each employment status

	All (%)	Stayers (%)	Emigrants (%)
Employment spells			
0	24.5	25.6	22.2
1	53.8	48.5	64.2
2	12.7	14.4	9.4
3	5.1	6.3	2.6
$\geqslant 4$	4.0	5.2	1.5
Unemployment spells			
0	80.6	75.0	91.6
1	12.6	15.9	6.1
2	4.4	5.8	1.7
$\geqslant 3$	2.4	3.3	0.6
Non-participation spells			
0	29.2	28.4	30.7
1	47.9	44.5	54.5
2	13.0	14.6	9.9
3	5.5	3.1	3.1
$\geqslant 4$	4.4	2.6	1.8

unemployment benefits, but about 70% experience a period without income. Emigrants more often do not experience any unemployment or no-income spells.

IV. Estimating wage growth of immigrants

Since the literature analysing immigrant earnings has established the need to separate wage growth from arrival cohort effects (Borjas, 1985; LaLonde and Topel, 1992) and the advantages of using longitudinal data to address unobserved heterogeneity, we first describe the classic models studied in the literature and use artificially constructed data similar to widely available data, enabling us to compare the estimates of that model with those where we control for out-migration selection and the endogeneity of employment and wages. The artificial data are all based on imposing availability restrictions on the rich and dynamic data we observe: For each migrant who entered between 1/1/1999 and 31/12/2007 we know the full migration history from entry until departure or until the end of the observation period (1/1/2008), whatever occurs first, and the full labour market history while they are in the country.

First, in section 'cross sectional data' we assume that we observe a survey conducted at 1/1/2008 containing all the recent migrants who are in the country at 1/1/2008. Assuming a survey implies that we only observe their current (monthly) income, employment status and all other relevant individual characteristics. We know the entry date and can derive from this the years-since-migration variable. Based on these data we obtain OLS estimates and estimates from a Heckman selection model that corrects for selective employment status. In section 'panel stock data', we follow the literature in using panel data and hence assume

that we observe for all these stayers in 1/1/2008, the monthly employment and income history (and the history of other relevant migrant characteristics). Based on these panel data we obtain the fixed effects (FE) and random effects (RE) estimates for the earnings profile of migrants.

Both individual labour market behaviour and out-migration are likely to depend on factors that also influence the wage growth. This renders the timing of employment status changes and out-migration endogenous in the wage growth analysis. In the next subsections we discuss dynamic models that explicitly model the endogeneity of the timing of individual labour market behaviour, out-migration or both. In section 'stock data with full labour market history', we describe a correlated multi-state model for the stayers with three labour market states, which takes into account the full labour market history and their interdependence. In section 'accounting for selective out-migration' we also include migrants who leave the host country before the end of the observation window, utilising the information on the timing of their departure. We formulate a model that allows for interdependence between the wage profile and out-migration, yet ignoring the (possible endogenous) labour market transitions. The results of this model will provide insights into the importance of accounting for selective out-migration in measuring the wage profile of immigrants. In section 'combining labour market history and selective out-migration', we describe our full model, which extends the correlated multi-state model with an absorbing living-abroad state to capture selective out-migration. This model fully utilises all the dynamic information available in the data. We examine all migrants as a homogenous group without differentiating between family and labour migrants since almost none of the previous studies distinguish between the two.

Cross sectional data

Data is often collected at fixed periods, e.g. one-off survey data. Our starting point is replicating the standard methodology used by previous cross-sectional studies. We assume that the entry date is available; thus, for a migrant who arrived mid-1999 and stayed in the country till the end of the observation window, we have nine observation points, each 1 January of each year, 2000–08. The standard Mincer wage equation is (considering only positive wages):

$$\ln W = \theta_0 + \sum_{h=1}^{8} \theta_h \mathrm{YSM}_h + \theta_x X + u_w, \tag{1}$$

where X are other control variables that influence the wage, namely gender, age, age square and industry sector.[11] We also control for both cohort and time effects, using year of arrival dummies as well as national unemployment rates in the Netherlands at the time of arrival. To capture the macro-economic effects that might affect wages, we include the national unemployment rate in 1/1/2008. Furthermore, arguing that wages are potentially correlated with out-migration, we control for the economic conditions of origin countries

[11] Unfortunately we do not have information on education. Lubotsky (2007) also does not control for individuals' educational attainment, arguing that this allows for measuring un-conditional differences in earnings.

and potential pull factors using log GDP and GDP per capita growth in the country of origin in 2008. We assume, as is standard in the literature, $u_w \sim \mathcal{N}(0, \sigma^2)$.

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

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

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

Panel stock data

Next we assume that the data are from a retrospective panel, based on the number of migrants still in the country at the end of the observation period. Pischke (1992) and Borjas (1989), for example, among others, used such longitudinal data. For these migrants we observe their full monthly employment and income history (as well as the history of other relevant migrant characteristics). Such data allow us to estimate both random effects (RE) and fixed effects (FE) panel models.

It is standard practice in the literature to estimate the wage growth of immigrants considering only those migrants who are still in the country at the end of the observation window which in our case is 1/1/2008. These immigrants have stayed in the country for a maximum of nine years $h = 1, \ldots, 8$, or 108 months. We use real wages, and estimate the following equation based on the (unbalanced) panel of monthly wage observations :

$$\ln W_{im} = \theta_i + \sum_{h=1}^{8} \theta_h \text{YSM}_h + \theta_x X_{im} + u_{im}, \tag{3}$$

where W_{im} is the real wage observed for individual i in month m, YSM_h are years-since-migration dummies- which depend on the individual arrival and observation month- and X_{im} are control variables that influence the wage. We control for demographics, monthly sector of employment, as well as the Dutch business cycle using monthly unemployment rate. We also control for pull factors using annual country of origin's economic conditions. We also account for cohort and period effects. For the cohort effect, we use both the unemployment rate in the Netherlands at the time of entry (measured on a quarter-year basis) and dummies for the year of arrival. The period effects are captured by year dummies.[12] Of course, when using a fixed effects model, the time constant control variables are excluded, e.g. gender and the year of arrival dummies. We assume that the errors are normally distributed, $u_{im} \sim \mathcal{N}(0, \sigma_u^2)$ and, when using a random effects model, $\theta_i \sim \mathcal{N}(0, \sigma_\theta^2)$.

[12] Note that the common identification problem of age-period-cohort is not an issue here because the period and cohort variables are not co-linear. First, the YSMs are not changing at the same time as the period and year of arrival dummies, because we observe the exact entry times. Second, the year of arrival dummies are not co-linear with the unemployment rate at entry, because the latter changes with each quarter.

Panel data have a distinct advantage, namely they allow controlling for several potential biases due to correlated individual effects. Such biases arise from the cohort effects, and the role of age at entry. Nonetheless, panel data cannot control for out-migration selection if attrition cannot be disentangled from out-migration. Panel attrition will bias the estimation results if the probability of leaving the sample is systematically linked to labour market status and wages.

Stock data with full labour market history

The wage growth of an immigrant strongly depends on the labour market performance of the migrant in the host country. Migrants who are employed all the time are more likely to secure higher wages than those who are unemployed or non-participatory for a certain period. It is very likely that the transitions of an individual migrant among these labour market states depend on both observed and unobserved factors, which also influence the wage growth. In this section, we diverge from the previous literature by showing the significance of the labour market history. As in the previous section we first restrict the analysis to 'stayers', namely the immigrants who remain in the host country until/beyond the end of our observation window. We formulate a correlated competing risk model (a correlated multivariate mixed proportional hazards model), expanded with a log-linear (correlated) wage equation. This employs the observed labour market history of these stayers. We call this model panel stock data labour-CCRM, or labour-CCRM for short. An immigrant potentially faces several labour market transitions during their stay in the host country and multiple durations in the different states. We consider three labour market states, together with a wage equation:

1. Employed in the host country (e);
2. Unemployed and receiving benefits in the host country (u);
3. Out of the labour market (includes both unemployed but not receiving benefits and non-labour market participants) in the host country (non-participating, n);

These states are mutually exclusive and exhaust all possible destinations. A migrant may leave a labour market state $j = \{e, u, n\}$ for any of the other labour market states, e.g. for an employed migrant ($j = e$) the destination states are unemployment and out of the labour market ($k = u$ or $k = n$). We explicitly model the timing of the labour market changes. We view the labour market behaviour of an individual migrant as a multivariate Mixed Proportional hazard, with individuals moving between the three labour market states. All hazard rates depend on the duration of the ongoing spell.

For each transition from j to k, say from employment to unemployment, we define the duration variables T_{jk} that describe the time since entry into the labour market state j and assume a mixed proportional hazard model for each transition rate:

$$\lambda_{jk}(t|\overline{X}_{jk}(t), V_{jk}) = \lambda_{0jk}(t) \exp\left(\beta'_{jk} X_{jk}(t) + V_{jk}\right) \tag{4}$$

where $\overline{X}_{jk}(t) = \{X_{jk}(s)|0 \leqslant s \leqslant t\}$ is the sample path of the observed characteristics up to time t. The unobserved heterogeneity V_{jk} also enters the hazard multiplicatively. We assume that the path of the observed characteristics is independent of the unobserved heterogeneity. We assume that the baseline duration dependence $\lambda_{0jk}(t)$ is piecewise constant on R

intervals,[13] i.e. $\lambda_{0jk}(t) = \sum_{r=1}^{R} e^{\alpha_{jkr}} I_r(t)$ with $I_r(t) = I(t_{r-1} \leqslant t < t_r)$ and $t_0 = 0$, $t_R = \infty$. Any duration dependence can be closely approximated by arbitrarily increasing the number of intervals. For identification, we assume that the baseline hazard is one in the first interval, i.e. $\alpha_{jk1} = 0$.[14]

The wage of a migrant is only observed while the migrant is employed and has a standard log-linear Mincerian relation that depends on the time spent in the country (YSM) and other (possibly varying with time in the host country) covariates

$$\ln W(t) = \theta_0 + \sum_{h=1}^{H} \theta_h \mathrm{YSM}_h(t) + \theta_x x(t) + \epsilon(t), \tag{5}$$

where, for a given migrant, the error term comprises two components, an independently normally distributed idiosyncratic component, $\eta(t)$, and a random individual-specific component, v_w

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

All labour market transition hazard rates as well as wages are assumed to depend on observed and unobserved characteristics. The various unobserved characteristics (random effects) are allowed to be correlated in an unrestricted fashion, implying that the parameters of the model can be estimated using the non-parametric maximum likelihood estimator (Heckman and Singer, 1984; Gaure, Røed and Zhang, 2007).

For all transitions from a particular labour market state (j), only the shortest duration $\tilde{T}_j = \min_k T_{jk}$ and the corresponding actual new labour market state is observed. The other durations are censored, in the sense that we only know that their realizations exceed $\tilde{T}_j$. If for individual i we observe M_{ijk} spells, for j to k transitions, at sojourn times $t_1, \ldots, t_M$, then the likelihood contribution of these M_{ijk} transitions (conditional on unobserved heterogeneity vector V) is:

$$L_{jk}(V) = \prod_{m=1}^{M_{ijk}} \lambda_{jk}(t_m | \overline{X}_{jk}(t_m), V_{jk})^{\delta_{mjk}} \exp\Big(- \sum_{g \neq j} \Lambda_{jg}(t_m | \overline{X}_{jg}(t_m), V_{jg}) \Big) \tag{6}$$

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

The likelihood contribution (conditional on unobserved heterogeneity vector V) from a sequence of wage observations over an employment spell is the product of the densities of each observed wage

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

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

[14] In principle we could also achieve identification by restricting the expected value of the unobserved heterogeneity to one.

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

We allow for possible interdependence between the wage equation and the transition rates by assuming a joint, correlated distribution for the individual-specific component in the wage equation and all the unobserved heterogeneity terms in the transition rates. For the sake of parsimony, we assume that each of the unobserved heterogeneity terms remains the same for recurrent durations of the same type, and we adopt a joint discrete distribution, i.e. V has the discrete support $(V_1,\ldots,V_M)$ and $p_m = \Pr(V = V_m)$.[15] It is important to note that the V_m's are vectors with e.g. $V_m = (V^m_{e,u}, V^m_{e,n}, V^m_{u,e}, V^m_{u,n}, V^m_{n,e}, V^m_{n,u}, V^m_w)'$ including all the possible unobserved heterogeneity factors of the labour transitions and the random component of the wage equation. Thus, we assume that there are M latent types of individuals in the population with type m having the unique pair V_m of 'location' points, consisting of the six unobserved heterogeneity terms in the transition rates and the one individual specific component of the wage equation, and comprising the fraction p_m of the population.

The complete likelihood function for each individual is the integral over the (discrete) distribution of individual specific components of the product of the likelihood contribution of the labour market transition in equation (6) and the likelihood contribution of the observed wage development in equation (7)

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

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

Accounting for selective out-migration

As has previously been highlighted in the literature, out-migration is a major concern for the estimation of immigrant wage growth. To gain insights into the importance of accounting for selective out-migration we first ignore labour market transitions. From our (day-exact) information on in- and out-migration times we can construct the time spent in the Netherlands, or censored time when the migrant is still in the country at the end of the observation window. We follow common practice in duration analysis and the conditional hazard of leaving the host country follows a mixed proportional hazard model, given by the product of the baseline hazard and the exponentiated function of observed time-varying characteristics X and unobserved characteristics V_m:

$$\lambda(t|\overline{X}(t), V_m) = \lambda_0(t) \exp\left(\beta' X(t) + V_m\right) \tag{9}$$

where $\lambda_0(t)$ represents the baseline hazard, namely the duration dependence of the out-migration rate common to all migrants.

To account for selective out-migration, we model the out-migration rate and the wage formation simultaneously, similar to the multivariate Mixed Proportional hazards model in the previous section. Again, the wage of a migrant is only observed while the migrant is employed and depends on the time spent in the country (YSM) and other (possibly varying

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

with time in the host country) covariates, as well as a (discrete) unobserved factor (see equation (5)). We adopt a joint discrete distribution for the unobserved heterogeneity of the out-migration rate and the wage equation.

Combining labour market history and selective out-migration

It is very likely that the labour market behaviour, the out-migration and the wage development are interdependent. In this section we extend the labour-CCRM model from section 'stock data with full labour market history' to include an additional state of living abroad (out-migration) and consider a four-state multistate model correlated competing risks model (full-CCRM): A multivariate correlated mixed proportional hazards model with states (e) employed; (u) unemployed; (n) non-participating and the new (absorbing, we ignore repeated immigration) state (m), living abroad (out-migration; i.e. left the host country).

Again these states are mutually exclusive and exhaust all possible destinations. A migrant may leave a labour market state $j = \{e, u, n\}$ for any of the other destination labour market states including out-migration. Just as in labour-CCRM described in section 'stock data with full labour market history', we assume that migrant behaviour can be described as a multivariate correlated Mixed Proportional Hazards model with individuals moving between the first three states and now with abroad as an additional absorbing state.[16]

Again, the wage of a migrant is only observed while the migrant is employed and depends on the time spent in the country (YSM) and other (possibly varying with time in the host country) covariates, as well as a (discrete) unobserved factor, as in equation (5). Just as before, we adopt a joint discrete distribution for the unobserved heterogeneity of the multistate transition process and the wage equation, now consisting of 10 individual specific components: the nine unobserved heterogeneity terms in the transition rates and the one individual specific random component in the wage equation.

Note that this model allows for a direct impact of the labour market status on the decision to migrate, as we formulate different return hazards by labour market status.

Identification

Our main aim in this paper is to address the endogeneity of the labour market status/ wages and out-migration when examining immigrant earnings growth. In the CCRM models, the years-since-migration indicators are endogenous, as is the included lagged labour status. This may raise identification issues, with respect to both the transition rates of (repeated) competing risks and the wage equation. For the mixed proportional hazard models (MPH), we use standard practice – dating back to Heckman and Singer (1984) – for the unobserved heterogeneity based on a discrete distribution and an unknown number of mass points in which the correlation structure is completely flexible. Under some regularity conditions, such models have been shown to be non-parametrically identified from single spell data (see Van den Berg, 2001). Gaure *et al.* (2007) provide Monte Carlo evidence that modelling

[16] One limitation is that we assume that unobserved heterogeneity is time invariant. But we model dynamic selection through a time varying probability of labour market status.

correlation between hazards with a discrete distribution works well in the context of a timing-of-events model. Horny and Picchio (2010) prove that the lagged dependencies are identified with a MPH structure.

The identification of our hazard rates is strengthened by the presence of repeated (labour market) spells (Honoré, 1993; Van den Berg, 2001; Bijwaard, 2014). Moreover, we also condition on time-varying explanatory variables (e.g. national unemployment rate, GDP per capita and GDP growth rate of the country of origin). Brinch (2007) shows that exogenous explanatory variables that exhibit significant variation over time improve identification.

Similar multivariate mixed proportional hazard rate models (MMPH) have been used in the literature, mainly for analysing labour market transitions. For example, Uhlendorff and Zimmermann (2014) use such a model to compare the labour market behaviour of natives and migrants.

Only a few researchers have also integrated the effect on wages within a correlated competing risk model with mixed proportional hazards, e.g. McCall and Chi (2008) and Gaure *et al.* (2012). Despite not providing a formal proof, Gaure *et al.* (2012) argue that non-parametric identification of the MMPH model with an endogenous wage equation is through the MMPH part of the model. In other words, given our methodology, our model is identified. Nonetheless, we also add exclusion restrictions to the wage equation by excluding the marital status of the migrant and the number of children from the wage equation, since being married and having children are likely to be correlated with labour market participation decisions but not directly with wages.[17] The correlation of all the unobserved terms in the labour market transitions and the out-migration transition with the unobserved wage factor is thus identified through the observed distribution of wages conditional upon the labour market and migration behaviour (and other included explanatory variables). In addition, migrants with multiple employment spells add to the identification.

V. Results

The covariates included in the models refer to demographic characteristics (gender, age at entry, marital status and number of children), the country of origin's GDP per capita and GDP growth rate, as well as the labour market sector.[18] We control for business cycle conditions by including the national unemployment rate, considering both the unemployment rate at the quarter of entry into the country and the time-varying monthly rate. The quarterly unemployment rate at entry captures the 'cohort effect' of migrants, while the current varying unemployment rate captures the impact of the business cycle on the intensity to leave. Interestingly, we find that the covariates matter as determinants of wages: females have lower wages compared to males. Age and age square display the usual concave relationship with wages. Those immigrants from high income countries tend to have higher

[17] Without these additional, but not necessary, exclusion restrictions the results are similar.

[18] Ideally we would have also liked to control for labour market conditions in all other countries, however, given the complexity of our models and the number of controls we have, we only control for push factors in the host country and conditions in the country of origin. The education level of the immigrants is also not available. We argue that the sector of employment together with the country of origin and the migration motive serve as a rough approximation of the skill level of the immigrant. Besides, our models account for unobserved individual characteristics that influence the labour market behaviour and the out-migration decision. These unobserved factors are likely to capture the influence of skill differences.

Table 3

Impact of years since migration (YSM) on log real wage of migrants

	Cross sectional		Panel stock data			Full history	
	OLS-employed	Heckman	RE†	FE†	Labour-CCRM	Migration	Full-CCRM
YSM1‡	0.131**	0.130**	0.137**	0.144**	0.130**	0.138**	0.122**
	(0.005)	(0.005)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
YSM2	0.189**	0.186**	0.210**	0.224**	0.211**	0.177**	0.192**
	(0.006)	(0.006)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
YSM3	0.212**	0.207**	0.247**	0.267**	0.252**	0.199**	0.231**
	(0.006)	(0.006)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)
YSM4	0.236**	0.232**	0.261**	0.288**	0.271**	0.206**	0.248**
	(0.007)	(0.007)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)
YSM5	0.243**	0.238**	0.265**	0.297**	0.276**	0.205**	0.250**
	(0.008)	(0.008)	0.002)	(0.003)	(0.003)	(0.003)	(0.003)
YSM6	0.267**	0.263**	0.262**	0.300**	0.275**	0.199**	0.251**
	(0.009)	(0.009)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)
YSM7	0.256**	0.251**	0.259**	0.302**	0.264**	0.179**	0.238**
	(0.010)	(0.010)	(0.003)	(0.004)	(0.004)	(0.005)	(0.004)
YSM8	0.244**	0.238**	0.284**	0.328**	0.237**	0.143**	0.208**
	(0.014)	(0.014)	(0.006)	(0.007)	(0.007)	(0.008)	(0.007)

Notes: In all models the following control variables are included: gender, age and age-squared, logarithm of the GDP of the country of origin and the GDP growth, national unemployment rate (at entry and during stay), entry year and sectoral dummies. In the migration and CCRM models piecewise constant duration dependence is assumed and marital status and number of children are included in the transition rates. In the CCRM-models, we also include repeated (un)employment or no income indicators. A third-order polynomial of the log-wage while employed is included in the migration transition and in the transitions from employment in the CCRM models.** $P < 0.01$.† RE: Random effects; FE: Fixed effects.‡ YSM1: Years-since-migration 1–2 year; YSM2: Years-since-migration 2–3 years; etc.

wages, and national unemployment rates tend to have negative impact on wages.[19] We also include cohort and period effects. In the fixed effects model, only time-varying variables are included. In the models including labour market history (CCRM-labour and full-CCRM), the individual's labour market history (indicators of previous (un)employment and non-participation experience) is also included and we assume a piecewise constant baseline hazard on six intervals (at three and six months and at one, two and five years and beyond five years for the hazard of leaving employment, as well as at three and six months, one, two and three years and beyond three years for the hazard of leaving unemployment and non-participation). In the models including the out-migration timing (migration and full-CCRM), we assume a piecewise constant baseline hazard on six intervals at six months, one year, 18 months, two and five years and beyond five years) for the out-migration hazard.[20]

Table 3 provides the estimated impact of years-since-migration on the immigrant earnings profile for all the described models. The estimates from the cross sectional OLS model seem to underestimate the earnings' profile compared to the panel models.[21] The

[19] The full table with estimated coefficients for the migration-model, the labour-CCRM, and the full-CCRM model can be found in Table B.1–B.3 in the Online Appendix B.

[20] We have tried different piecewise-constant hazard specifications and found no significant change in the estimated YSM-effects.

[21] Table 1 shows the wage growth with no controls, while the OLS estimates in Table 3 provide the estimates controlling for observables and the full-CCRM estimates control for both observables and unobservables characteristics.

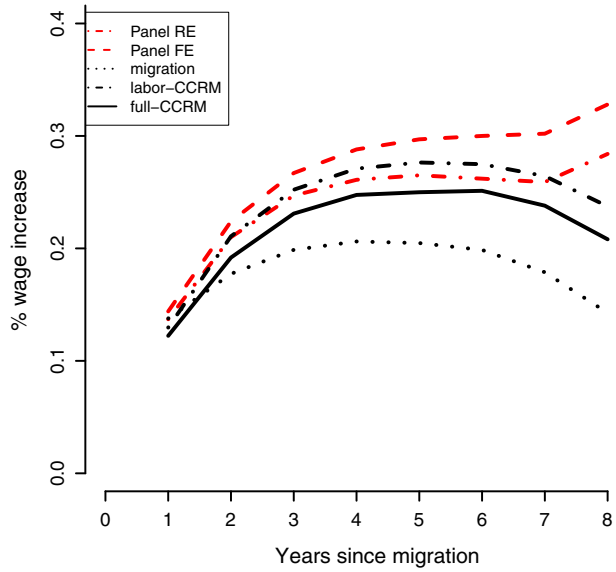


Figure 1. Immigrants' wage growth by years since migration

Heckman corrected OLS estimates are also slightly lower than the OLS estimates. The estimates in the fourth column indicate that the use of fixed effect panel data suggests a higher impact of YSM on wage growth than the other methods. In fact, estimates from panel data seem to provide higher wage growth compared to those obtained from simple cross sectional data. This suggests that not controlling for previous labour market status and current macro-economic conditions reduces the wage growth. Furthermore, the OLS underestimates immigrants' wage growth unlike studies done on the US, possibly because in the case of the Netherlands, as we can see in Table 1, the average immigrant who leaves tends to have had higher wages, which suggests positive selection for out-migration. This is the reverse of the US, where immigrants with low earnings are disproportionately more likely to emigrate from the United States over time; i.e. out-migrants are negatively selected (see e.g. Lubotsky, 2007). Accounting only for selective out-migration (column 6) underestimates the wage growth since it ignores labour market transitions and therefore periods of non-employment, while accounting only for the labour market behaviour (column 5) overestimates the wage growth compared to the full correlated competing risks model (column 7) since it does not take into account the selectivity of out-migration and in particular that successful migrants out-migrate. This suggests that not accounting for out-migration and labour market status would bias the wage growth of immigrants in the case of the Netherlands. The difference between the results from the full-CCRM and the migration model are larger than the difference between the results from the full-CCRM and the labour-CCRM, implying that accounting for selective labour market transitions is more important than accounting for selective out-migration. Figure 1 depicts these estimated immigrant earnings profiles for the panel stock data and the CCRM models.

From the relative difference between the estimates of each model with respect to our preferred full-CCRM model (see Table B.4 in Online Appendix B) we draw some additional

conclusions: The cross sectional estimates which control for observables only show that the bias is not homogeneous in one direction. This suggests that selection on unobservables is not time-constant. Also, the full-CCRM estimates potentially indicate two opposite types of biases: first, an upward bias resulting from controlling for labour market transition, which would lead to higher wages; and a downward bias resulting from the out-migration of successful immigrants, resulting in lower wages. Below, we examine whether those patterns exist for different types of migrants.

VI. Distinguishing by migration motive

The economics literature on migration has typically focused on the behaviour of labour migrants, despite the fact that many migrants have non-labour-related motives to migrate. In both Europe and the US, family reunification and family formation are the cornerstone of immigration policies; 34% of recent migrants to the Netherlands, for example, are such family migrants. These migrants have less attachment to the host labour market. Jasso and Rosenzweig (1995) and Aydemir (2011) have shown that skilled immigrants achieve relatively greater success on the labour market compared with family reunification immigrants. Moreover, these family migrants are also generally less mobile than labour migrants (Bijwaard, 2010). Nonetheless, only a few studies have explored the behaviour of family migrants in detail. Jasso and Rosenzweig (1995) focus on family migrants to the US, Aydemir (2011) to Canada, Miller (1999) to Australia, Rodriguez Planas and Vegas (2011) to Spain and Bijwaard and van Doeselaar (2014) the Netherlands.

In order to understand better the selectivity into out-migration and the effect of the labour market status, we differentiate between labour and family migrants, who are commonly indistinguishable in the census or commonly-used panel data. Table B.5 in Online Appendix B provides the characteristics of migrants by migration type. One striking albeit unsurprising difference is that almost three-quarters of all family migrants are females, compared to less than one third among labour migrants. Half of the labour migrants are from EU15, and 14% are from the new EU (A8 countries). On the other hand, 62% of family migrants are from LDCs, of whom 20% are from Turkey and 18% are from Morocco. Therefore, we additionally categorise the LDC family migrants by these two countries. At the end of the observation window (1/1/2008), 84% of the labour migrants are employed compared to 42% among family migrants, although labour migrants are likely to be more recent (or have fewer YSM).

Unsurprisingly, only 5% of family migrants were employed at the time of entry, although 42% became employed at the end of our observation period. On the other hand, all labour migrants were employed upon arrival, but only 84% were employed at the end of our observation period. Thus, it is clear that the labour market transitions are different between those two types of migrants and such transitions are likely to have different effects on both wage growth and out-migration selection.

Figure 2 depicts the Kaplan–Meier estimates of the survival probability of remaining in the Netherlands by migration motive. It clearly shows that the out-migration behaviour strongly differs between labour and family migrants. For both migration motives, income plays an important role in explaining out-migration. When we run a simple Cox proportional hazard model on the time spent in the Netherlands with income (including non-wage income

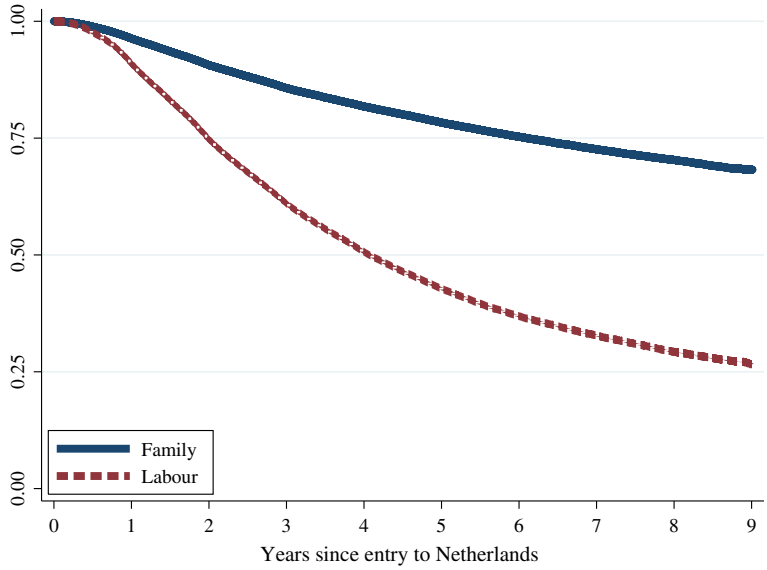


Figure 2. Kaplan–Meier estimates of survival probability of immigrants to remain in the Netherlands

and zero income, in category dummies) as the only explanatory variable, whereby we find a clear U-shaped income effect on out-migration, similar to Bijwaard and Wahba (2014) who only use initial income, with no income and the highest income groups leaving the soonest.

Another interesting difference between labour and family migrants is in their income at entry and the evolution of their wage earnings as depicted in Table B.6 in Online Appendix B. The income at entry is clearly much lower for family migrants, even after conditioning for a positive wage. The increase in the average income for family migrants with YSM is due to their increase in paid employment. For labour migrants, those who leave have a higher wage compared to those who stay. Indeed, examining the earning growth profile of all labour migrants masks the higher earnings of those who leave: those labour migrants who are continuously employed (both emigrants and stayers) experience the highest wage growth. This is not the case for family migrants who do not experience the same growth in their wages even if they are continuously employed. For family migrants, it is clear that the labour market status is very important. This confirms that it is problematic to pool these two types of migrants, given that they behave differently in the labour market and in terms of out-migration.

Table B.7 in Online Appendix B depicts the employment dynamics for the labour and family migrants separately. A very clear distinction between the labour market behaviour of labour migrants and family migrants exists. By definition, all labour migrants have employment experience, whereas only 53% of the family migrants do. For family migrants the difference in labour market behaviour between stayers and emigrants is also larger. The family migrants who remain in the country are more mobile on the labour market than those who leave. Labour migrants who remain are also slightly more mobile on the labour market, except that they less often experience non-participation spells.

Findings by migration motive

The results for the models that account for selection for the two migration motives are depicted in Figure 3 below (see Table B.8 in Online Appendix B). It should be emphasised that the estimates for the whole sample (both types of immigrants) shown in Figure 1 are not a simple weighted average of the labour and family migrants. The selection processes for both types of immigrants may be different and thus when combining them (into the whole sample) we might lose part of the selection because we assume that it is similar for all immigrants. Furthermore, the wage of family migrants increases faster than the wages of labour migrants. Of course, many family migrants are not employed at or soon after arrival and tend to begin employment at a lower wage, whereby they experience a greater opportunity for wages to grow.

Ignoring the interdependence of the labour market processes leads to a biased estimation of the wage growth for both types of migrants, albeit the bias is greater for the family migrants compared to labour migrants. For labour migrants, controlling only for out-migration is more important in early years of migration whilst labour market endogeneity seems to matter more later on, suggesting that there is negative selection in the early years and positive selection in later years. On the other hand, for family migrants, ignoring labour market endogeneity underestimates the wage growth given the high labour market transition they experience, whereas ignoring selective out-migration overestimates their wage growth.

From the relative difference between the estimates of the labour-CCRM and the migration model with respect to the full-CCRM model, as shown in Figure 3 (see Table B.9 in Online Appendix B) we can derive some interesting insights: For labour migrants, the full-CCRM estimates uncover two selection processes that work in the same direction – both pulling down the wages. For family migrants, the full-CCRM estimates mask two opposing selection processes: an upward bias resulting from controlling for selective labour market transition, which would lead to higher wages, and a downward bias resulting from the selective out-migration. More importantly, even when examining one type of migrants,

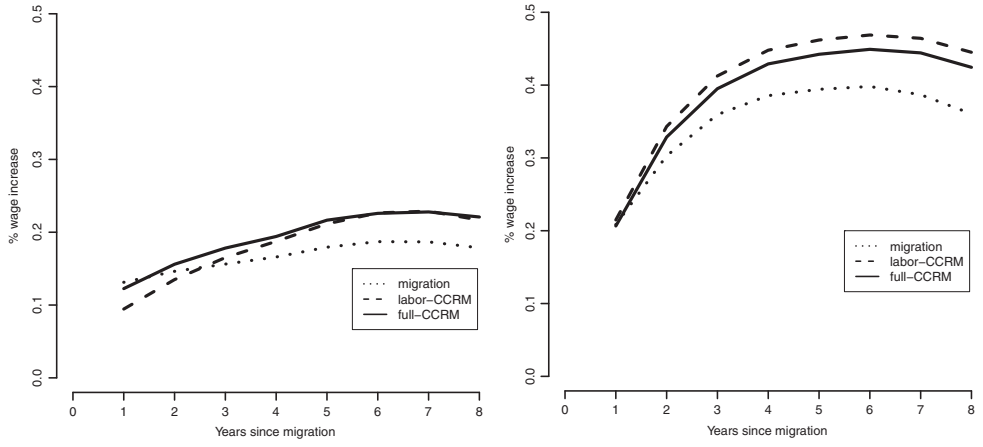


Figure 3. Wage growth by years since migration, by migration motive (labour migrants left, family migrants right)

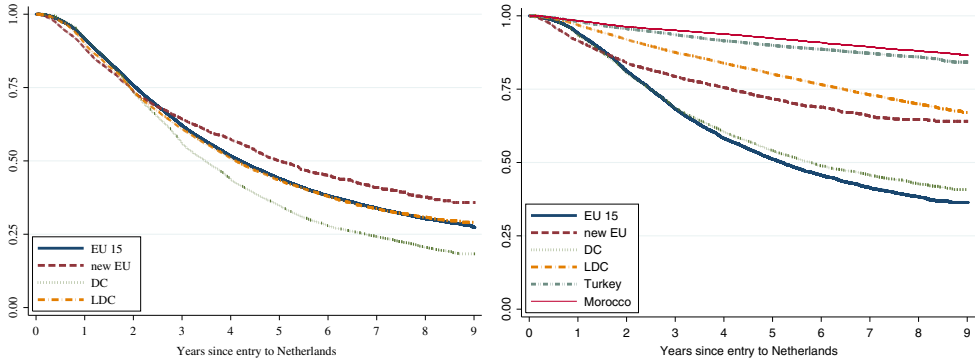
the results show that selectivity of out-migration varies over time; i.e. that the biases are time variant. Overall, our estimates underscore the importance of differentiating between types of immigrants given the difference in labour supply and out-migration probabilities influencing their earnings profiles.

Findings by region of origin and migration motive

In this section, we distinguish between the region/country of origin given the differences in migration restrictions faced by immigrants. Accordingly, we distinguish between four main groups: immigrants from the EU15 who can move freely in the Dutch labour market; immigrants from the new EU; immigrants from other developed countries (DC) and immigrants from developing countries (LDC). For family migrants we further split the LDC-migrants into migrants from Turkey, Morocco and the remaining LDCs. Figure 4 depicts the Kaplan–Meier estimates of the survival probability of remaining in the Netherlands of labour migrants and family migrants from the main countries of origin.

For both migration motives we see clear differences among the regions of origin vis-à-vis the likelihood to remain in the country. Labour migrants leave much sooner overall, and immigrants from non-EU DCs leave sooner for both migration motives. Family migrants from Turkey or Morocco hardly ever leave and for family migrants from the remaining LDCs and from new EU-countries, departure is also rare.

Again it should be emphasised that the selection processes for immigrants from different regions may be different and when combining them (into the whole sample by migration motive) the estimated wage profiles are not a weighted sample of the wage profiles by region. Interestingly, when we distinguish amongst labour migrants by region/country of origin, the full-CCRM estimates (Figure 5, left panel) highlight that labour migrants from the DCs experience the smallest wage growth compared to those from the new EU, EU and LDCs (see Tables B.10 and B.11 in Online Appendix B for the full estimates). This is because labour migrants from DCs tend to have the highest initial wage at€6,069 compared to€2,906 for EU labour migrants,€1,740 for new EU labour migrants and€3,131 for LDCs labour migrants. The right panel of Figure 5 shows that the returns to YSM for family migrants increase over time and tend to be lower than those of labour migrants, except



(a) Labour migrants (b) Family migrants

Figure 4. Kaplan–Meier estimates of survival probability of remaining in the Netherlands, for main countries of origin by migration motive (labour migrants left, family migrants right)

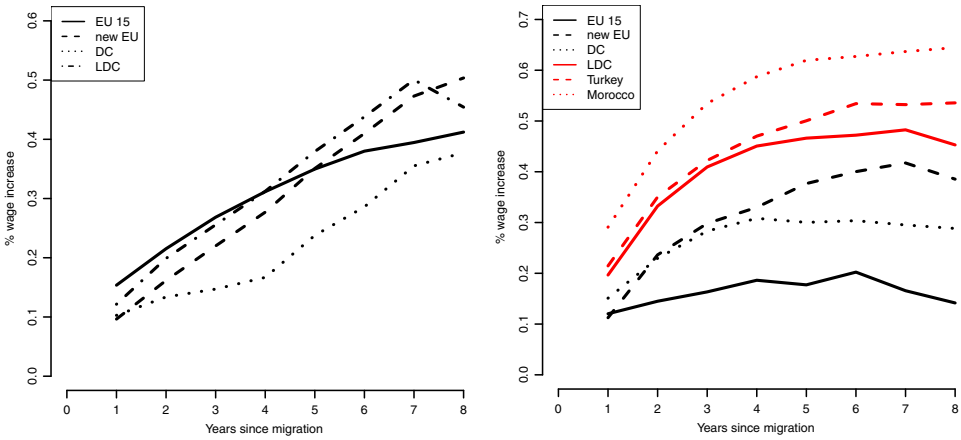


Figure 5. Wage growth by years since migration (full-CCRM), by region and migration motive (labour migrants left, family migrants right)

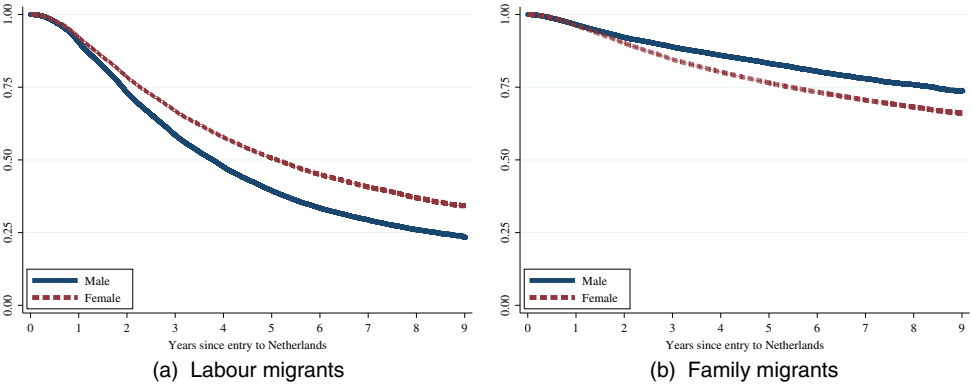


Figure 6. Kaplan–Meier estimates of survival probability of remaining in the Netherlands, by gender and migration motive (labour migrants left, family migrants right)

for family migrants from Morocco who experience the largest wage growth, as they tend to have the lowest initial wage (€786 compared to €936 for family migrants from Turkey and €1,182 for family migrants from other LDCs). Indeed, immigrants from the same country/region experience different wage growth depending on whether they are labour or family migrants, i.e. the region/country of origin matters as much as the migrant type for the wage growth. This confirms that out-migration and labour market selectivity are host-country- as well as origin-country-specific.

Findings by gender and migration motive

Although we control for a number of labour market decisions, including labour force participation, it is important to examine the wage growth of men and women separately. First, the literature tends to focus on male migrants only. Secondly, since females are disproportionally represented among family migrants, it seems also particularly relevant to

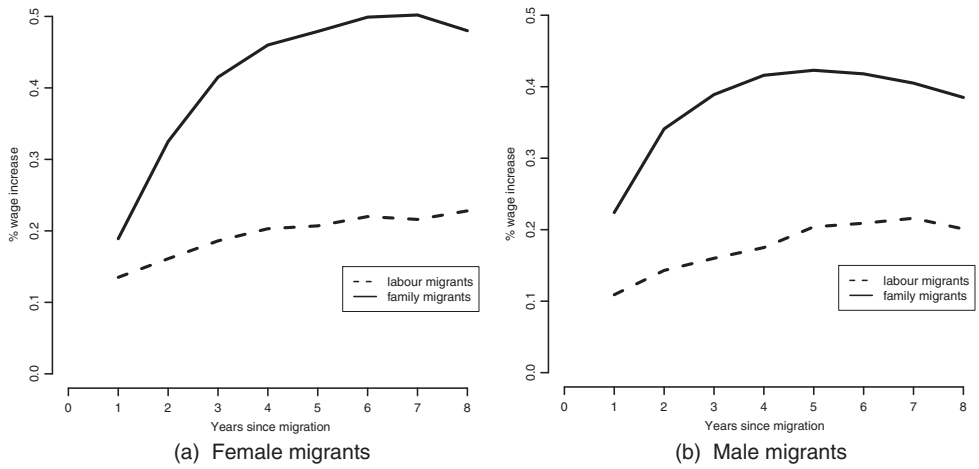


Figure 7. Wage growth by years since migration (full-CCRM), male vs female migrants by migration motive

identify whether controlling for gender and migration motives biases (due to not controlling for selective labour market and/or selective out-migration) reveals a large differential. Figure 6 depicts the Kaplan–Meier estimates of the survival probability of remaining in the Netherlands for labour migrants and for family migrants by gender, clearly showing that out-migration differs substantially by gender: female family migrants leave the Netherlands sooner, whereas among labour migrants it is the men who leave sooner.

A very clear distinction between the labour market behaviour of males and female family migrants also exists in our data.[22] Female family migrants have much less employment experience and more unemployment and non-participation spells. For labour migrants the gender difference in labour marker behaviour is much less pronounced. Still males have more stable employment and less unemployment and non-participation experience. For female labour migrants the influence of migration behaviour on the labour market experience seems larger.

The estimation results for the models (see Tables B.12 and B.13 in Online Appendix B) that account for labour market and migration selection for both genders, separately by migration motives are depicted in Figure 7. Again it should be noted that the estimates for the whole sample (both genders) shown in Figure 3 are not a simple weighted average of the male and female migrants. The selection processes for both genders may be different and thus when combining them (into the whole sample) we might lose part of the selection because we assume that it is similar for all immigrants. Furthermore, the wage of female family migrants increases faster than the wages of male family migrants. For labour migrants the gender difference in wage growth is small.

Overall, our results suggest that migration motives are more important for the wage growth than just gender differences, as labour migrants behave differently from family migrants both in term of labour market participation and out-migration. Hence our results highlight the importance of joint modelling out-migration and wage endogeneity when estimating wage earnings of immigrants, as those two decisions operate differently for different types of immigrants.

[22] Detailed information is available upon request.

Robustness checks

In this sub-section, we check the robustness of our results by estimating earnings profiles of immigrants by year of arrival to control for cohort effects. We focus on the oldest cohort of immigrants; i.e. those who arrived in 1999. Interestingly, examining the descriptive evolution of average earnings of the 1999 entry cohort (Table B.14 in Online Appendix B) clearly reveals that emigrants (those who out-migrated) have experienced the largest growth in their wages, particularly for those who were continuously employed. At the same time, emigrants show a more depressed wage growth independent of employment status. As previously shown, emigrants' earnings are higher than stayers'.

Interestingly, analysis of the estimates from our three main models (migration, labour-CCRM, and full-CCRM: Table B.15 in Online Appendix B) confirms that controlling only for out-migration underestimates the wage growth, and controlling only for labour market interdependence overestimates the wage growth. Based on the 1999 cohort the estimated wage growth is higher than before, since those are the oldest cohorts. However, the findings confirm the conclusion from the whole sample that the role of out-migration selection and of labour market selectivity differs by migration motive (Table B.16 in Appendix B).

Furthermore, we have also checked the sensitivity of our results to using the initial wage at arrival (first observed wage) as the reference wage – instead of the wage at month three – and all our results are robust. Not surprisingly, by using wage upon arrival as the reference wage, we obtain a noticeable jump in the wage and therefore in the wage growth between year one and year two. We also checked the robustness of our findings by using wages at month six as the reference wage, and again find that our results are robust and similar to those using wages at month three. Finally we also checked the robustness of our results against using nominal wages as opposed to real wages. Again all our findings are qualitatively unchanged.[23]

VII. Conclusion

The performance of immigrants in the labour market – and particularly their earnings – is an important indicator of their contribution to the economy. However, examining the earnings profiles of immigrants over their migration cycle is challenging. This paper examines immigrant wage growth, taking into account the out-migration selection and the labour market selection using administrative data from the Netherlands. More importantly, we address an additional challenge that has hitherto been overlooked in this literature, namely the notion that out-migration might be correlated with the labour market status and earnings of immigrants. We also distinguish between labour and family migrants, whose labour market and out-migration behaviours are different and yet are typically not disentangled. A further distinction by region of origin and gender also shows substantial differences in the wage growth.

We first apply standard cross sectional data methods assuming our data is only collected retrospectively at the end of the observation period. Next we apply panel data methods using random and fixed effects – as previously undertaken in the literature – by using similar

[23] The results from the robustness checks using a different reference wage – wages at six months as well as nominal wages – are not included here but are available upon request from the authors.

(constructed) data to provide us with benchmark results. Next, we estimate a model with selection into out-migration, albeit ignoring the labour market transitions. We subsequently estimate a correlated multi-state model (labour-CCRM), which takes the endogeneity of wages and labour market status into account, but ignores the out-migration. Our final model combines these last two models into a multi-state model (full-CCRM) that takes into account the out-migration selectivity as well as the endogeneity of the labour market status.

We find that classic panel models lead to overestimates of immigrants' wage growth. We also find that controlling only for out-migration selectivity or controlling only for labour market selectivity also provides biased estimates. Accounting for selective labour market transitions seems as important as accounting for selective out-migration. Controlling for out-migration selectivity only would underestimate the wage profiles of migrants as it ignores the labour market transitions and periods of non-employment, while controlling for the labour market selectivity would overestimate the wage profile, as it doesn't account for those who out-migrate because of being successful in the labour market. Indeed, our results highlight the importance of taking into account the endogeneity of the labour supply and out-migration when estimating immigrant earnings profiles.

More importantly, our findings show that different selections are important for different types of migrants and that selectivity is changing over time. For labour migrants, controlling only for out-migration is more important in early years of migration whilst labour market endogeneity seems to matter more later on, suggesting that there is negative selection in the early years and positive selection in later years. On the other hand, for family migrants, ignoring labour market endogeneity underestimates the wage growth given the high labour market transition they experience – whereas ignoring selective out-migration would overestimate their wage growth. Hence, accounting for labour market endogeneity seems as important as accounting for selective out-migration, particularly for family migrants. Overall, our findings underscore that family and labour migrants behave differently, and hence experience different forms of selection. Thus, not distinguishing between migrant types when estimating immigrant wage profiles is problematic.

Final Manuscript Received: November 2018

References

Abbring, J. H. and van den Berg, G. J. (2003). 'The non-parametric identification of treatment effects in duration models', *Econometrica*, Vol. 71, pp. 1491–1517.

Abramitzky, R., Boustan, L. P. and Eriksson, K. (2014). 'A nation of immigrants: assimilation and economic outcomes in the age of mass migration', *Journal of Political Economy*, Vol. 122, pp. 467–506.

Anteco, H., Kuhn, P. and Trejo, S. J. (2006). 'Assimilation via prices or quantities? Sources of immigrant earnings growth in Australia, Canada, and the United States', *Journal of Human Resources*, Vol. 41, pp. 821–840.

Aydemir, A. (2011). 'Immigrant selection and short-term labour market outcomes by visa category', *Journal of Population Economics*, Vol. 24, pp. 451–475.

Barth, E., Bratsberg, B. and Raaum, O. (2004). 'Identifying earnings assimilation of immigrants under changing macroeconomic conditions', *Scandinavian Journal of Economics*, Vol. 106, pp. 1–22.

Biavaschi, C. (2016). 'Recovering the counterfactual wage distribution with selective return migration', *Labour Economics*. Vol. 38, pp. 59–80.

Bijwaard, G. E. (2010). 'Immigrant migration dynamics model for The Netherlands', *Journal of Population Economics*, Vol. 23, pp. 1213–1247.
Bijwaard, G. E. (2014). 'Multistate event history analysis with frailty', *Demographic Research*, Vol. 30, pp. 1591–1620.
Bijwaard, G. E. and van Doeselaar, S. (2014). 'The impact of changes in the marital status on return-migration of family migrants', *Journal of Population Economics*, Vol. 27, pp. 961–997.
Bijwaard, G. E. and Wahba, J. (2014). 'Do high-income or low-income immigrants leave faster?', *Journal of Development Economics*, Vol. 108, pp. 54–68.
Bijwaard, G. E. and Wang, Q. (2016). 'Return migration of foreign students', *European Journal of Population*, Vol. 32, pp. 31–54.
Bijwaard, G. E., Schluter, C. and Wahba, J. (2014). 'The impact of labour market dynamics on the return-migration of immigrants', *Review of Economics and Statistics*, Vol. 96, pp. 483–494.
Borjas, G. J. (1985). 'Assimilation, changes in cohort quality, and the earnings of immigrants', *Journal of Labour Economics*, Vol. 3, pp. 463–489.
Borjas, G. J. (1989). 'Immigrant and emigrant earnings: a longitudinal study', *Economic Inquiry*, Vol. 27, pp. 21–37.
Borjas, G. J. and Bratsberg, B. (1996). 'Who leaves? The outmigration of the foreignborn', *The Review of Economics and Statistics*, Vol. 78, pp. 165–176.
Brinch, C. N. (2007). 'Nonparametric identification of the mixed hazards model with time-varying covariates', *Econometric Theory*, Vol. 23, pp. 349–354.
Chiswick, B. R. (1978). 'The effect of Americanisation on the earnings of foreign-born men', *Journal of Political Economy*, Vol. 86, pp. 897–921.
Dustmann, C and Weiss, Y. (2007) 'Return migration: theory and empirical evidence', *British Journal of Industrial Relations* Vol. 45, pp. 236–256.
Dustmann, C. (1995). 'Return migration: The European experience', *Economic Policy*, Vol. 22, pp. 214–250.
Dustmann, C. and Görlach, J.-S. (2015). 'Selective outmigration and the estimation of immigrants earnings profiles', in Chiswick B. R. and Miller P. W. (eds), *Handbook on the Economics of International Migration*, Vol. 1A, chapter 10, Amsterdam: North–Holland, pp. 489–533.
Edin, P.-A., Lalonde, R. J. and Åslund, O. (2000). 'Emigration of immigrants and measures of immigrant assimilation: evidence from Sweden', *Swedish Economic Review*, Vol. 7, pp. 163–204.
Galor, O. and Stark, O. (1991). 'The probability of return migration, migrants' work effort, and migrants' performance', *Journal of Development Economics*, Vol. 35, pp. 399–405.
Gaure, S., Røed, K. and Zhang, T. (2007). 'Time and causality: a Monte Carlo assessment of the timing-of-events approach', *Journal of Econometrics*, Vol. 141, pp. 1159–1195.
Gaure, S., Røed, K. and Westlie, L. (2012). 'Job search incentives and job match quality', *Labour Economics*, Vol. 19, pp. 438–450.
Heckman, J. J. and Singer, B. (1984). 'A method for minimising the impact of distributional assumptions in econometric models for duration data', *Econometrica*, Vol. 52, pp. 271–320.
Honoré, B. E. (1993). 'Identification results for duration models with multiple spells', *Review of Economic Studies*, Vol. 61, pp. 241–246.
Horny, G. and Picchio, M. (2010). 'Identification of lagged duration dependence in multiple spell competing risks models', *Economics Letters*, Vol. 106, pp. 241–243.
Hu, W. Y. (2000). 'Immigrant earnings assimilation: estimates from longitudinal data', *American Economic Review*, Vol. 90, pp. 368–372.
Jasso, G. and Rosenzweig, M. R. (1982). 'Estimating the emigration rates of legal immigrants using administrative and survey data: the 1971 cohort of immigrants to the US', *Demography*, Vol. 19, pp. 279–290.
Jasso, G. and Rosenzweig, M. R. (1995). 'Do immigrants screened for skills do better than family reunification immigrants?', *International Migration Review*, Vol. 29, pp. 85–111.
Jensen, P. and Pedersen, P. J. (2007). 'To stay or not to stay? Out–migration of immigrants from Denmark', *International Migration*, Vol. 45, pp. 87–113.
LaLonde, R. J. and Topel, R. H. (1992). 'The assimilation of immigrants in the U. S. labor market', in Borjas G. J. and Freeman R. B. (eds), *Immigration and the Workforce: Economic Consequences for the United States and Source Areas*, Chicago, IL: University of Chicago Press, pp. 67–92.

Lubotsky, D. (2007). 'Chutes or ladders? A longitudinal analysis of immigrant earnings', *Journal of Political Economy*, Vol. 115, pp. 820–867.
McCall, B. and Chi, W. (2008). 'Unemployment insurance, unemployment durations and re-employment wages', *Economics Letters*, Vol. 99, pp. 115–118.
Miller, P. (1999). 'Immigration policy and immigrant quality: the Australian points system', *American Economic Review*, Vol. 89, pp. 192–197.
Pischke, J. S. (1992). *Assimilation and the Earnings of Guest Workers in Germany*, Discussion Paper No. 92-17, ZEW, Mannheim.
Rodriguez Planas, N. and Vegas, R. (2011). 'Moroccans' assimilation in Spain: family-based versus labour-based migration', *Middle East Development Journal*, Vol. 3, pp. 119–139.
Roy, A. (1951). 'Some thoughs on the distribution of earnings', *Oxford Economic Papers*, Vol. 3, pp. 135–146.
Uhlendorff, A. and Zimmermann, K. F. (2014). 'Unemployment dynamics among migrants and natives', *Economica*, Vol. 81, pp. 348–367.
Van den Berg, G. J. (2001). 'Duration models: specification, identification, and multiple duration', in Heckman J. and Leamer E. (eds), *Handbook of Econometrics*, Vol. V, chapter 55, Amsterdam: North–Holland, pp. 3381–3460.
Van Ours, J. and Veenman, J. (2005). 'The Netherlands; old emigrants – young immigrant country', in Zimmermann K. (ed), *European Migration: What Do We Know?* Oxford: Oxford University Press, pp. 173–196.
Wahba, J. (2014). 'Return migration and economic development', in Lucas R. E. B. (ed), *International Handbook on Migration and Economic Development*, Cheltenham UK and Northampton, MA, USA: Edward Elgar Publishing Ltd, pp. 327–349.
Zorlu, A. and Hartog, J. (2002). 'Migration and immigrants: the case of The Netherlands', in Rotte R. and Stein P. (eds), *Migration Policy and the Economy: International Experiences*, München: Hans Seidel Stiftung, pp. 119–140.

Supporting Information

Additional supporting information may be found in the online version of this article:

Appendix A. Intstitutional backround: Dutch immigration rules 1999–2007.
Appendix B. Additional Tables.