NORFACE MIGRATION

NORFACE MIGRATION Discussion Paper No. 2011-7

The Impact of Labour Market Dynamics on the Return–Migration of Immigrants

Govert E. Bijwaard, Christian Schluter and Jackline Wahba

www.norface-migration.org

The Impact of Labour Market Dynamics on the Return–Migration of Immigrants*

Govert E. Bijwaard †
Netherlands Interdisciplinary Demographic Institute (NIDI)

Christian Schluter ‡
DEFI, Aix-Marseille II and University of Southampton

Jackline Wahba §
University of Southampton

May 2011

Abstract

Using administrative panel data on the entire population of new labour immigrants to The Netherlands, we estimate the causal effects of labour dynamics on their return decisions. Specifically, the roles of unemployment and re-employment spells on immigration durations are examined. The endogeneity of labour market outcomes and the return migration decision, if ignored, confounds the causal effect. This empirical challenge is addressed using the "timing-of-events" method. We estimate the model separately for distinct immigrant groups, and find that, overall, unemployment spells shorten immigration durations, while re-employment spells delay returns for all but one group. The magnitude of the causal effect differ across groups.

Keywords: temporary migration, durations, timing of event method, labour market dynamics.

JEL Codes: J61, J64, C41

*Financial support from the NORFACE research programme on Migration in Europe – Social, Economic, Cultural and Policy Dynamics is gratefully acknowledged. We thank Statistics Netherlands, and in particular Ruben van Gaalen, for data support.

†Corresponding author: `bijwaard@nidi.nl`

‡`c.schluter@soton.ac.uk`

§`j.wahba@soton.ac.uk`

1 Introduction

The labour market performance of immigrants in the host country has received ample attention in the empirical literature. Neglected, however, is the question as to what extent this labour market performance affects the decisions of migrants to return to their source country. In particular, what is the effect of adverse or positive labour market events such as the occurrence of unemployment spells and re-employment spells ?[1] How does the effect vary by the duration of the labour market spell ? The failure of the empirical literature to ask these questions let alone to furnish convincing answers arises from a combination of methodological challenges and severe limitations of the data usually encountered in migration analysis.

We address these novel questions using a unique administrative panel for the entire population of recent immigrants to the Netherlands covering the years 1999-2007. These data characteristics - large size, repeated and accurate measurement - are unique in migration analysis and enable us to examine durations reliably. The usual data situation is one of small samples, possibly subject to selectivity and attrition issues, extracted from surveys of respondent who provide recall data.

This Dutch immigrant register is based on the legal requirement for immigrants to register with the authorities upon arrival.[2] Several other official registers are linked by Statistics Netherlands to this immigrant register, such as the social benefit and the income register (used by the tax authorities). Sojourn times in labour market states are thus accurately recorded - they are day exact. Consequently, no data based on individual recall has to be used, nor do we employ less precise interval estimation techniques for durations. Moreover, the usual concerns about measurement error are less acute.

Another important feature of our data is the report in the immigrant register of the immigration motive, also recorded by the authorities upon arrival of the migrant. This enables us to focus explicitly and exclusively on 94,270 labour immigrants. The immigration motive is usually latent in standard datasets, and different behavioural patterns of labour and non-labour migrants would confound the empirical analysis. Indeed, the descriptive analysis of Bijwaard (2010) reveals the substantial extent of non-labour immigration to The Netherlands (75% of immigration is found to be non-labour related), as well as the substantial heterogeneity by immigration motive. Finally, the size of our data allows us to estimate our models separately for distinct immigrant groups defined in terms of their labour mobility at entry by immigration laws. In particular, we consider immigrants from sending countries in the EU15 ('old Europe'), the new EU (the majority of which are Poles and arrived after 2004); the countries outside Europe are grouped into developed (DCs) and less developed (LDCs) sending countries.

Turning to the methodological problem of estimating the *causal* effect of the labour market dynamics on the return decision of the immigrant, we need to control for unobserved correlated heterogeneity in the labour market and migration processes. Otherwise the resulting endogeneity confounds the causal impact (below we actually quantify the resulting bias and reveal it to be substantial). This end is achieved by using the "timing-of-events" method (Abbring and van den Berg (2003)). In particular, employment, unemployment and migration durations are modelled as mixed proportional hazards which incorporate correlated

[1] Existing models of the return decision yield conflicting empirical predictions, and are silent about the effects of such shocks. The two principal opposing paradigms are theories of optimal migration durations based on preference for source country consumption, and theories of mistaken expectations and immediate failure on the labour market.

[2] They are also required to de-register upon leaving. Non-compliers are removed from the register by administrators. We take this 'administrative removal' into account in our modelling, see Section 2.3

unobservables. The model for the migration duration permits the sojourn times in the various labour market states to have causal effects. These are then estimated non-parametrically using piecewise constant functions.

Having thus overcome the principal empirical challenges by using appropriate estimation techniques on rich population data on labour immigrants, this paper enlarges the evidence base for policy makers. As immigration has become a core public concern in most developed economies, policy makers seek to manage immigrant stocks. Understanding the link between the labour market and migration processes is fundamental to this end. In particular, quantifying the (possibly non-linear) effects of unemployment durations on the return migration decision is relevant to current debates about the financial costs, in terms of the state's social welfare bill, of "failed" immigrants. Such debates also usually ignore that the labour market fortunes of these immigrants can be reversed; hence we also consider the effects of re-employment.

We find that, unconditionally, both unemployment and return migration are prevalent: between 29% and 37% of labour immigrants experience unemployment spells, and 47% leave the host country during the observation window of 1999 to 2007. Turning to the causal effects, overall, unemployment spells shorten immigration durations, while re-employment spells delay returns for all but one group. The magnitude of the causal effects differ across groups. We further quantify the causal impact of labour market dynamics in terms of migration durations in several experiments, focussing on the duration and timing of unemployment spells, and, in a counterfactual analysis, the effect of improved immigrant "quality". We find that the unemployment durations have a substantial effect, while the effect of differences in timing and "quality" are relatively smaller.

The outline of the paper is as follows. We consider briefly the related literature next. The econometric model is set out in detail in Section 2. In particular, we specify the labour market and the migration processes, and elucidate the role of unobservable heterogeneity in Section 2.1. The causal effect is identified by the argument of Section 2.2. Estimation proceeds by maximising the likelihood which is spelt out in Section 2.4. The data are described in Section 3, and the empirical results are presented in Section 4. The last section concludes. A data appendix provides additional information.

1.1 The related literature

Several theories have been advanced to explain the return decision of migrants, sometimes with conflicting empirical predictions. Since none of these focus on the effects of adverse or positive labour market shocks during the migration spell, we do not intend to test these individual theories; their concern is simply different. We briefly summarise some prominent examples from the theoretical literature on the return migration motive for the sake of completeness.

According to one class of theories, return migration is planned and part of an optimal strategy to maximise life-time utility. The return of the migrant in e.g. Galor and Stark (1991) is due to higher preference for consumption in the own country relative to that in the host country. Dustmann and Weiss (2007) show that 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. In this set up, preference for the home country leads to return even though it is not necessarily economically advantageous to do so. Another optimal strategy explains return migration as a result of target savings where individuals migrate temporarily for a period of time where wages are higher so that they can accumulate savings overseas (Dustmann (2003)). Alternatively, individuals may migrate temporarily to acquire skills that are highly rewarded in the source country (Dustmann (1997)).

A fundamentally different mechanism is based on mistaken expectations about, and immediate failure on the host country's labour market, leading to an 'unplanned' return (Borjas and Bratsberg (1996)). If this mechanism is at work, return migration is expected to take place relatively soon after arrival in the host country.

Several empirical investigations have attempted to identify the return motive albeit with mixed conclusions. For example , Yang (2004) finds evidence in support of the life cycle explanation. On the other hand Gibson and McKenzie (2009) conclude that the decision to return is strongly linked to family and lifestyle reasons rather than to the income opportunities in different countries. To repeat, we do not intend to test these theories of the return migration motive. Rather, we consider the effect labour market shocks experienced during the immigration.

Closer to our concern is Kirdar (2007) who considers, like us, the effect of unemployment spells on the return decision. However, his analysis does not address the confounding selection bias, and is based on a small panel for Germany which exhibits the standard data characteristics discussed above.

Return migration from The Netherlands has been examined before by Bijwaard (2009, 2010). However, the focus of these descriptive analyses is different from ours. Bijwaard (2010) does not focus on labour migrants, nor uses data on the labour market status of the migrants. He shows that the migration dynamics of recent migrants is substantial, and heavily depends on the migration motive and the country of birth. Labour migrants and students are much more prone to leave than family migrants. Bijwaard (2009) considers the correlation between migration decisions and labour market status transitions. But he does not estimate the causal effect of the labour market process on the return decision, nor does he take the endogeneity of labour market transitions and migration decisions into account.

2 The econometric model

We seek to determine the *causal* effect of labour market dynamics on the return migration intensity of immigrants, so the outcome random variable of interest is the time spent in the Netherlands. The empirical challenge arises from the potential correlation between the labour market process and the migration process which, if present and ignored, would confound the causal effect. We address this endogeneity issue by allowing for correlated unobservable heterogeneity.

The observational units are labour immigrants in the host country (the Netherlands). We consider first the random time T since first entry a certain event takes place: Let T_m denote the time the immigrant leaves the host country in order to return to the sending country, T_e the time an employment spell ends in the host country, and T_u the time an unemployment spell ends. We also define the associated time-varying indicators: the indicator $I_u(t)$ takes value one if the migrant is unemployed at time t, and $I_e(t)$ indicates that the immigrant is employed again after a period of unemployment. Finally, define by $\delta_e(t)$ and $\delta_u(t)$ the sojourn time in the employment and unemployment state. As the migrant is either employed or unemployed, the labour market process is alternating. We thus have a correlated migration and labour market status process (employed or unemployed), with three possible transitions: (i) unemployment to employment; $T_{e1}, T_{e2}, \ldots$, (ii) employment to unemployment; $T_{u1}, T_{u2}, \ldots$, and (iii) return migration (leaving the Netherlands); $T_{m1}, T_{m2}, \ldots$, which thus allows for multiple migrations and returns.

The timing of events and our definitions are illustrated in Figure 1. We depict the labour market and migration durations of two migrants. In accordance with our data definitions

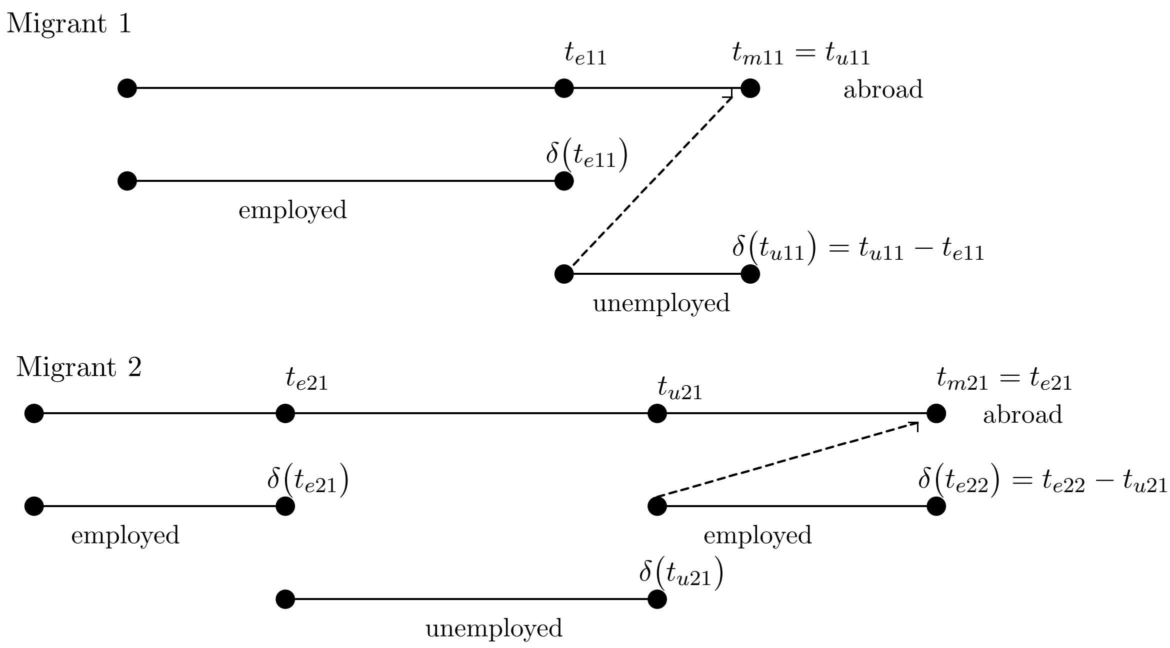


Figure 1: Migration and labour market dynamics

of Section 3, the migrants are employed at the moment they enter the country. Migrant 1 arrives after Migrant 2. The length of Migrant 1's (first) employment spell is $\delta\big(t_{e11}\big) = t_{e11}$. He remains in the country unemployed until time t_{m11}. His unemployment spell is thus of duration $\delta\big(t_{u11}\big)= t_{u11} - t_{e11}$. The unemployment spell is terminated at the moment he leaves the host country at time t_{m11}. Migrant 2 stays for a longer period in the country, $t_{m21} > t_{m11}$, and undergoes a different labour market experience. His first employment spell has duration $\delta\big(t_{e21}\big) = t_{e21}$. After an unemployment spell of length $\delta\big(t_{u21}\big) = t_{u21} - t_{e21}$ he becomes employed again. This second employment spell is terminated when he leaves the host country at time t_{m21}, and has duration $\delta\big(t_{e22}\big)= t_{e22} - t_{u21}$. The last labour market spell for each migrant is always censored. While Migrant 1 experiences an adverse labour market shock (unemployment), Migrant 2 experiences a positive shock (re-employment). We seek to determine the effect of such shocks, both in terms of their incidences and their durations, on the duration of the migration spell.

We allow T_m, $\delta_e(t)$ and $\delta_u(t)$ to be correlated through unobservable heterogeneity terms and through a possible direct effect of labour market dynamics on the migration intensity. The latter is the causal effect we seek to estimate. To be precise, we express the distributions of the random variables in terms of their hazard rates $\theta_e\big(\delta_e(t)|x_e(t),\, v_e\big), \theta_u\big(\delta_u(t)|x_u(t), v_u\big)$ and $\theta_m\big(t|t_e, t_u, x_m(t), v_m\big)$. $v = (v_e, v_u, v_m)$ collects the possibly correlated unobservable heterogeneity terms, which are distributed according to some distribution function G. Time varying characteristics in state $k = \{u, e, m\}$ are denoted by $x_k(t)$. Conditional on x and v the distributions of T_u and T_e are independent. The migration incidence and the labour market changes are characterised by the dates they occur, and we are interested in the effect of the realisation of unemployment, t_u, and re-employment, t_u on the distribution of T_m.

In particular, we assume that the conditional hazards follow mixed proportional hazard models, given by products of baseline hazards (measuring duration dependence) and functions of observed time-varying characteristics x and unobserved characteristics v:

$$\theta_u\big(\delta_u(t)\big|x_u(t),v_u\big) \;=\; v_u\lambda_u\big(\delta_u(t)\big)\exp\Big(x_u(t)\beta_x^u\Big) \tag{1}$$

$$\theta_e\big(\delta_e(t)\big|x_e(t),v_e\big) \;=\; v_e\lambda_e\big(\delta_e(t)\big)\exp\Big(x_e(t)\beta_x^e\Big), \tag{2}$$

and

$$\theta_m(t|t_u,t_e,x_m(t),z(t),v_m) = v_m\lambda_m(t)\exp\Bigg(x_m(t)\beta_x^m + I_u(t)\Big\{\gamma_u+\alpha_u\big(\delta_u(t)\big)+z_u(t)\phi_u\Big\} + I_r(t)\Big\{\gamma_e+\alpha_e\big(\delta_e(t)\big)+z_e(t)\phi_e\Big\}\Bigg). \tag{3}$$

λ_k is the baseline hazard in state k, α_k are piecewise constant functions, and the covariates z in the (return) migration hazard are a subset of the time-varying characteristics of the migrants x. α_u is the causal effect of unemployment, and α_e is the causal effect of re-employment on the migration hazard. By employing piece-wise constant functions, we allow these effects to exhibit duration dependence.

2.1 Endogeneity: Confounding unobservable heterogeneity

We briefly discuss the important endogeneity issue. It is well known that, due to dynamic sorting effects, the distribution of v_u among those who become unemployed at t_u will differ from its population distribution. In particular, individuals with high v_e will tend to enter unemployment earlier than individuals with low v_e. If v_e and v_m are dependent, then the distribution v_m for unemployed migrants at a given time in the country will differ from the distribution of v_m for migrants still employed. Similarly, if v_m and v_u are not independent, then the distribution of v_m among unemployed migrants will differ its population distribution. Therefore, one cannot infer the causal effect of unemployment on the return-migration from a comparison of the realised durations of those who became unemployed at t_u with the rest of the population, because one would then mix the causal effect of unemployment on the duration with the difference in the distribution of v_m between these migrants. In this case $I_u(t)$ and $I_e(t)$ will be endogenous, and T_u, T_e and T_m should be modelled jointly to account for dependence of the unobserved heterogeneity terms. Therefore, we allow v_u, v_e and v_m to be correlated.

For the sake of parsimoniousness, we assume that each of the unobserved heterogeneity terms remains the same for recurrent durations of the same type, and we adopt a two-factor loading model with two independent fundamental factors W_1 and W_2, both having a discrete distribution on $(-1, 1)$ with $p_j = Pr(W_j = 1)$. This implies that

$$v_k = \exp(\alpha_{k1}W_1 + \alpha_{k2}W_2) \tag{4}$$

with $k = \{u, e, m\}$. Let $W = (W_1, W_2)'$, $v = (v_e, v_u, v_m)'$ and A be the matrix of factor loadings with rows $A_k = (\alpha_{k1}, \alpha_{k2})$. Then the variance-covariance matrix of the unobserved heterogeneity terms is given by $\text{Var}\big(\ln(v)\big) = A\text{Var}(W)A'$.[3]

2.2 Identification of causal effects: The timing-of-events method

The "timing-of-events" method of Abbring and Van den Berg (2003) makes the causal effect of employment dynamics identifiable by the functions $\gamma_u + \alpha_u\big(\delta_u(t)\big) + z_u(t)\phi_u$ and

[3]One additional restriction is needed for identification. We let $\alpha_{m2} = 0$.

$\gamma_e + \alpha_e(\delta_e(t)) + z_e(t)\phi_e$ where α_k is estimable non-parametrically. To achieve this identification we need to invoke their "no anticipation"assumption. In particular, we assume that the realisation of t_u (unemployment) only affects the intensity $\theta_m(t|t_e, t_u, x_m(t), v_m)$ for $t > t_u$, and the realisation of t_e (re-employment) only affects θ_m for $t > t_e(> t_u)$. This no-anticipation assumption rules out that migrants, knowing they will become unemployed (re-employed), will act upon this by leaving the country before they actually become unemployed (re-employed). Note however, that this does not require migrants to have no knowledge of the magnitude of the effect of a labour market change, nor to have no knowledge of the precise timing of when this labour market change will occur. It only requires that migrants do not modify their migration behaviour before the labour market change.

The no-anticipation assumption might fail for a small number of migrants who are on temporary contracts, and who anticipate their departure in advance. But this potential bias from anticipation should be small since even for these migrants the migration behaviour is not fixed in advance as they may renegotiate contract renewal or find another job that allows them to stay in the country. Moreover, the time span between the moment at which the anticipation occurs and the moment at which the actual labour market change is relatively short compared to the duration of stay in the country. This is another reason why the induced bias should be small. See Abbring and van den Berg (2003) for further discussion of the restrictiveness of this basic identification assumption.

2.3 Administrative removal

Some migrants do not officially inform the authorities that they are about to leave the host country. The return migration of the non-compliers is coded as an "administrative removal" by the authorities, once these have concluded that the immigrant has left the municipality, and has not registered with, and does not show up in the files of another municipality (this is further discussed below in data Section 3).

We address this as follows. We assume that the events "administrative removal" and "zero income at last observed time" imply that the migrant has left *before* the date the administrative removal is recorded, and *after* the last date of any observed change in the observed characteristics (e.g. labour market status, housing and marital status). Such limited information is equivalent to *interval-censored* data. For interval-censored data the exact end of a duration is unknown, but it is known that the duration ended in some period of time. If a migrant is administratively removed at duration t_a and the last observed change for this migrant occurred at duration $t_1 < t_a$, the contribution to the likelihood (of the out-migration) of this migrant is the probability of survival till t_1 times the probability that the migrant left the country between t_1 and t_a. The latter is equal to the survival from t_1 until t_a given survival. Consequently, administrative removal has no effect on the employed part of the likelihood function.

Let a_{ik} indicate whether the k^{th} emigration of migrant i was due to an administrative removal ($a_{ik} = 1$). For an administratively removed migrant we introduce two different event dates: t^a_{ik} is the administrative removal date of the k^{th} emigration of migrant i and $t^1_{ik} < t^a_{ik}$ is the date of the last recorded change in any of the characteristics of migrant i before t^a_{ik}.

2.4 Likelihood function

We have data for $i = 1, \ldots, n$ immigrants entering the Netherlands in our observation window. Let K_{ie}, K_{iu} and K_{im} be the number of observed employment, unemployment and migration spells of individual i. Note that for some migrants $K_{iu} = 0$ (e.g. a migrants who remains employed). We have three indicators $\Delta^e_{ik}, \Delta^e_{ik}$ and Δ^m_{ik} denoting that k^{th} employment/unemployment or migration spell is uncensored. Thus the likelihood contribution of migrant i conditional on the unobserved heterogeneity $v = (v_e, v_u, v_m)$ is, in the light of the preceding discussions:

$$\begin{aligned} L_i(v) &= \prod_{k=1}^{K_{iu}} \Bigg\{ \left[\theta_u\big(\delta_u(t_{ik})\big|\cdot, v_u\big)^{\Delta^u_{ik}} \exp\Big(-\int_0^{\delta_u(t_{ik})} \theta_u(\tau|\cdot, v_u)\, d\tau\Big)\right]^{(1-a_{ik})} \\ &\quad \cdot \left[\exp\Big(-\int_0^{\delta_u(t^1_{ik})} \theta_u(\tau|\cdot, v_u)\, d\tau\Big) - \exp\Big(-\int_0^{\delta_u(t^a_{ik})} \theta_u(\tau|\cdot, v_u)\, d\tau\Big)\right]^{a_{ik}} \Bigg\}^{I_u(t^-_{ik})} \\ &\quad \times \prod_{j=1}^{K_{ie}} \left[\theta_e\big(\delta_e(t_{ij})\big|\cdot, v_e\big)^{\Delta^e_{ij}} \exp\Big(-\int_0^{\delta_e(t_{ij})} \theta_e(\tau|\cdot, v_e)\, d\tau\Big)\right]^{I_e(t^-_{ij})} \\ &\quad \times \prod_{l=1}^{K_{im}} \left[\theta_m(t_{il}|\cdot, v_m)^{\Delta^m_{il}} \exp\Big(-\int_0^{t_{il}} \theta_m(\tau|\cdot, v_m)\, d\tau\Big)\right]^{(1-a_{il})} \\ &\quad \cdot \left[\exp\Big(-\int_0^{t^1_{il}} \theta_m(\tau|\cdot, v_m)\, d\tau\Big) - \exp\Big(-\int_0^{t^a_{il}} \theta_m(\tau|\cdot, v_m)\, d\tau\Big)\right]^{a_{il}} \end{aligned} \tag{5}$$

This likelihood naturally separates unemployment, employment, and migration spells, and for each spell allows for censoring and administrative removal. To simplify notation, we have suppressed the dependence on observed characteristics in the hazard rates. $I_u(t^-_{ik})$ indicates that the migrant is unemployed just before t_{ik} and similarly for $I_e(t^-_{il})$. When $K_{iu} = 0$ the relevant term becomes 1. Note that the last, and only the last, labour market spell is censored. This is either because the migrant is still in the country at the end of the observation period or because the migrant has left the country.

Integrating out the unobserved heterogeneity distribution we obtain the likelihood function

$$L = \prod_{i=1}^{n} \int\int\int L_i(v)\, dG(v_e, v_u, v_m) \tag{6}$$

where $G(v_e, v_u, v_m)$ is the joint distribution of the unobserved heterogeneity terms implied by the discussion of v_k given by equation (4).

3 Administrative panel data on the population of immigrants to The Netherlands

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

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

This immigration register is linked by Statistics Netherlands to the Municipal Register of Population (Gemeentelijke Basisadministratie, GBA) and to their Social Statistical Database (SSD). The GBA contains basic demographic characteristics of the migrants, such as age, gender, marital status and country of origin. From the SSD we have information (on a monthly basis) on the labour market position, income, industry sector, housing and household situation. Based on the income source Statistics Netherlands distinguishes nine labour market categories: employed, self-employed, unemployment benefits, disability benefits, social security benefits, other benefits, pensions, students and non-participating (no income). Note that most non-EU immigrants entering during our observation window do not qualify for social benefits, as eligibility requires sufficiently long employment or residence durations. We combine the first two labour market categories to define the individual's employment status. All the other categories are aggregated into the unemployed category.

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

To summarise the principal advantages of our data compared to conventional datasets used in the literature, we have a large panel of the entire population of labour immigrants; to be exact, we observe 94,270 individuals, 124,075 employment spells, and 59,248 unemployment spells. Other migration types, usually latent, are excluded and do not confound the empirical analysis which focuses on the effect of labour market dynamics. Income levels and labour market states are accurately recorded in the administrative data (as they are used by the authorities for tax and benefit purposes), and the start of the migration spell is recorded exactly. Moreover, the size of this labour immigrant population allows us to estimate our model separately for distinct migrant groups, rather than conducting a restrictive pooled analysis. In particular, we distinguish between migrants according to their initial labour mobility, and thus estimate separate models for migrants from sending countries in the EU15 ('old Europe'), the new EU (the majority of which are Poles and arrived after 2004), and the countries outside Europe are grouped into developed (DCs) and less developed (LDCs) sending countries. In the Data Appendix, we define these groupings precisely, and disaggregate these by country of birth.

Table 1: Descriptive dynamics

	EU15	new EU	non-EU DCs	non-EU LDCs
	Migration dynamics			
Mean spell per migrant				
Immigration	1.045	1.036	1.023	1.030
Emigration	0.533	0.300	0.601	0.424
Stayer	48.2%	70.9%	40.9%	58.6%
Once emigrated	50.5%	28.2%	58.3%	40.5%
>1 emigrated	1.4%	0.9%	0.9%	1.0%
In NL at end	51.2%	73.7%	42.3%	60.6%
	Length of stay at return migration			
<6 month	3.8%	12.6%	2.3%	5.9%
$6-12$ months	11.0%	19.6%	10.1%	13.3%
$12-18$ months	13.3%	16.8%	12.2%	12.9%
$18-24$ months	12.8%	13.8%	15.0%	13.2%
$24-60$ months	46.6%	30.1%	49.1%	43.8%
>5 year	12.5%	7.1%	11.3%	10.9%
Average [months]	33.2	24.5	33.2	31.1
	Labour market dynamics			
Mean spell per migrant				
Employment	1.398	1.272	1.115	1.253
Unemployment	0.747	0.497	0.418	0.529
Always employed	49.1%	62.9%	64.0%	62.1%
Once unemployed	36.5%	28.5%	31.9%	28.9%
>1 unemployed	14.4%	8.6%	4.1%	9.1%
Never re-employed	75.3%	80.9%	91.2%	84.1%
Once re-employed	16.0%	13.8%	7.2%	10.5%
>1 re-employed	8.7%	5.3%	1.7%	5.4%

3.1 Summary Statistics: Labour immigrants

We proceed to discuss summary statistics for our data relating to the dynamics of migration and of labour market events. The Data Appendix considers other aspects of the data.

In Table 1 we consider the incidence of return migration, and conditional on returning the duration of the stay in the Netherlands. Note that the group of 'stayers' includes permanent immigrants, and temporary migrants who have not yet returned. Hence immigrants from the new EU, having arrived predominantly in the second half of our observation window, are expected to exhibit a high proportion of censored migration spells. This is borne out in the data, since the share of stayers from the new EU is 71% whereas for other immigrants the range is between 41% and 59%. Relatedly, the durations of their completed spells are shorter. However, a large share of new EU movers (12.6%) leave the Netherlands after less than 6 months, which is considerably larger than for other immigrant groups. These differences highlight already the importance of a analysis disaggregated by sending countries.

Immigrants from the EU15 are more (less) likely to stay than migrants from (less) developed countries outside the EU, but more likely to be repeat migrants which reflects their unimpeded labour mobility. Conditional on returning, the distribution of completed durations look fairly similar for these three groups, as do the average duration. Turning to the unconditional distribution of the immigration duration, Figure 2 depicts the Kaplan Meier estimates of the survival probabilities by immigrant group. All groups look very similar for durations up to 24 months. One explanation for the differences at longer durations are the lower (higher) staying incidences for non-EU DCs (new EU). Relative to immigrants from the EU15, fewer immigrants from the latter group stay for longer durations. Overall, both Table 1 and Figure 2 highlight the importance of the temporary nature of labour migration. Across all immigrant groups, a substantial proportion leave the Netherlands eventually, and many do so within 24 months.

Turning to the labour market dynamics, Table 1 reveals that migrants from the EU15, relative to the other groups, experience greater labour market volatility: during the observation window they experience a higher incidence of unemployment spells (the mean spell is 0.7), more employment spells (1.4) and more than one unemployment spell (14.4%), and the share of the 'always employed' is smaller (49%).[4]

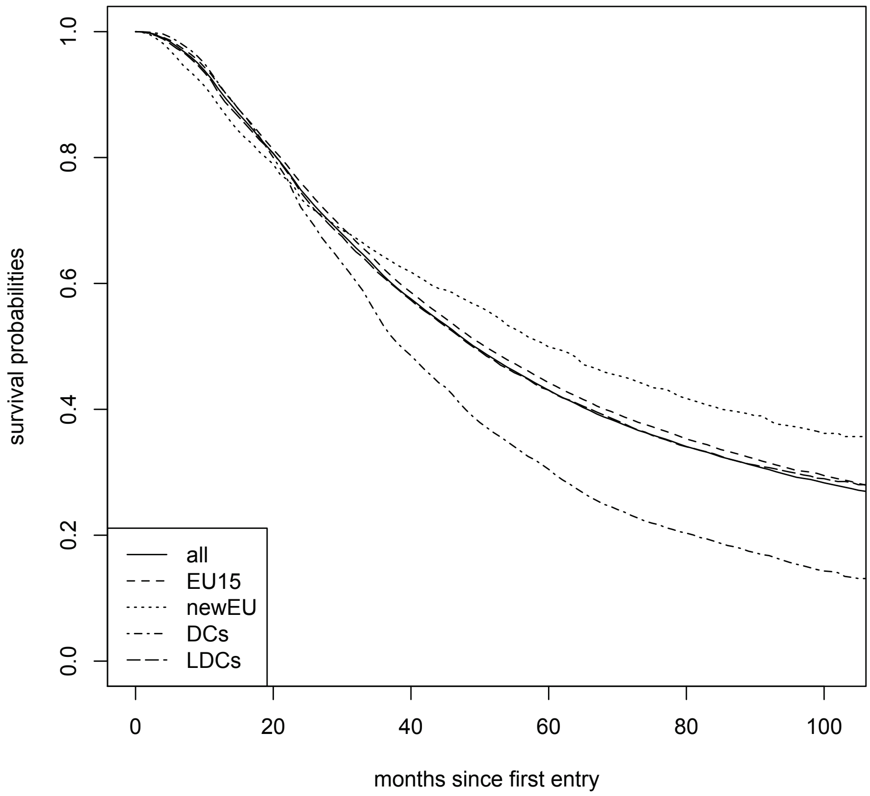
Since we seek to estimate the effects of negative and positive individual labour market shocks on the migration durations below, we now consider the immigrants by their labour market status prior to their departure from the host country. Hence Tables 2 and 3 condition on leaving the host country, whereas Table 1 considered the unconditional labour market dynamics.

In Table 2 we condition on being unemployed at the time of the return migration. In line with the results of Table 1, immigrants from the EU15 have a higher incidence of unemployment at the time of their departure (54%), a higher incidence of repeated unemployment (21%), and are more likely on average to experience longer unemployment durations (15 months). By contrast immigrants from DCs outside Europe have, compared to Europeans, lower incidences of unemployment (41%) and of repeated unemployment (8%), while their preceding employment spells were longer on average (20 months).

In Table 3 we consider immigrants who, after a period of unemployment, have found a job and subsequently leave. Hence this group has a volatile labour market experience (employment, followed by unemployment, followed by re-employment), but the last labour market spell is a 'positive' one. Unsurprisingly, the incidence of such labour market histories is low, ranging between 3 and 10%. Although non-European immigrants from DCs exhibit

[4]Note that migrants who are always employed, of course, enter the likelihood as censored observations, and are included in the estimation.

Figure 2: Kaplan Meier estimates of survival probabilities



the lowest incidence (3%), the durations of the last two labour market spells look fairly similar across all groups, except for the newEU immigrants who experience typically shorter durations.

Table 2: Descriptive statistics: Unemployed immigrants who leave.

	EU15	new EU	non-EU DCs	non-EU LDCs
Unemployed at emigration	53.8%	44.0%	40.7%	48.4%
Repeated unemployment	21.2%	15.0%	8.3%	15.9%
Mean # unemployment spells	1.32	1.21	1.10	1.24
	Current unemployment duration			
<3 month	18.6%	27.0%	23.3%	20.3%
$3-6$ months	16.0%	22.2%	17.2%	17.2%
$6-12$ months	22.7%	22.5%	19.4%	22.3%
>1 year	42.7%	28.3%	40.1%	40.2%
Average (months)	15.0	10.8	14.5	13.7
	Preceding employment duration			
<3 month	16.5%	17.8%	6.5%	12.0%
$3-6$ months	15.2%	17.8%	10.8%	11.9%
$6-12$ months	45.4%	38.8%	60.0%	53.6%
>1 year	22.9%	25.6%	22.7%	22.5%
Average (months)	15.5	13.2	20.0	17.8

Table 3: Descriptive statistics: Re-employed immigrants who leave.

	EU15	new EU	non-EU DCs	non-EU LDCs
Re-employed at emigration	7.5%	9.9%	2.8%	4.8%
repeated re-employment	32.0%	30.6%	12.3%	35.4%
Mean # re-employment spells	1.39	1.45	1.16	1.57
	Current (re-)employment duration			
<3 month	18.2%	26.9%	16.4%	17.0%
$3-6$ months	16.7%	21.8%	15.9%	17.3%
$6-12$ months	23.8%	22.8%	26.2%	23.9%
>1 year	41.3%	28.5%	41.5%	41.8%
Average (months)	14.8	10.5	13.9	14.3
	Preceding unemployment duration			
<1 month	18.4%	22.3%	11.3%	21.0%
$1-2$ months	18.1%	18.6%	17.9%	13.3%
$2-3$ months	17.2%	19.7%	17.4%	19.9%
$3-6$ months	20.5%	19.5%	19.1%	22.4%
$6-12$ months	17.1%	13.0%	25.6%	15.6%
>1 year	8.7%	6.9%	8.7%	7.8%
Average (months)	5.1	4.3	6.2	4.9

4 Results

We consider three versions of our model given by equations (1) to (3). Recall that the causal impacts of labour market dynamics on return migration hazards are given by $\gamma_k + \alpha_k\big(\delta_k(t)\big) + z_k(t)\phi_k$ with $k \in \{u, e\}$. In Model 1 the causal effects are assumed constant across time and migrants and thus given by γ_u and γ_e (with $\alpha_k \equiv \phi_k \equiv 0$). In Model 2 the covariate effect is still forced to zero, but the causal effects of sojourn times are allowed to exhibit duration dependence. This is implemented non-parametrically by modelling the sequence $\{\alpha_k\}$ as piece-wise constant functions. In Model 3 we further allow the effect to be heterogeneous across migrants in terms of z_u and z_e which measure demographics and previous labour market history. For the sake of expositional clarity, we present separately the effect of the unemployment and the re-employment spells, although it is clear from equation (6) that these are estimated simultaneously. For the sake of brevity, we do not discuss the coefficients of the covariates x_k which are only of secondary importance.[5] First we discuss the coefficients in the context of the return migration hazard. To facilitate the interpretation of the coefficients, we consider in Section 4.3 several illustrative examples which focus on the survival probabilities.

[5]The estimates are, of course, available from the authors. The covariates include extensive measures of demographics, a non-parametric function of income, housing descriptors, sector dummies, and controls for macro effects.

Table 4: The estimated causal effect of becoming unemployed on return migration hazards

	Model 1				Model 2				Model 3			
	EU 15	new EU	non-EU		EU 15	new EU	non-EU		EU 15	new EU	non-EU	
			DCs	LDCs			DCs	LDCs			DCs	LDCs
Constant effect [γ_u]	0.638**	0.670**	0.464*	0.667**								
	(0.076)	(0.165)	(0.185)	(0.193)								
Duration dependence [α_u]:												
(0-3 months)					0.633**	0.586**	0.663**	0.641**	0.834**	1.169**	0.403*	0.672**
					(0.078)	(0.171)	(0.178)	(0.197)	(0.085)	(0.206)	(0.192)	(0.205)
(3-6 months)					0.728**	0.877**	0.564**	0.797**	0.963**	1.401**	0.323	0.795**
					(0.080)	(0.175)	(0.181)	(0.202)	(0.087)	(0.209)	(0.194)	(0.208)
(6-12 months)					0.650**	0.657**	0.068	0.777**	0.915**	1.133**	-0.158	0.740**
					(0.081)	(0.180)	(0.183)	(0.207)	(0.088)	(0.212)	(0.196)	(0.209)
($>$ 1 year)					0.662**	0.681**	-0.224	1.035**	1.000**	1.025**	-0.433^{*}	0.991**
					(0.086)	(0.194)	(0.181)	(0.223)	(0.093)	(0.216)	(0.196)	(0.220)
Labour market history [ϕ_u]:												
Repeated unemployment									-0.179^{**}	-0.388^{*}	-0.017	-0.432^{**}
									(0.056)	(0.196)	(0.218)	(0.136)
Order of unemployment spell									-0.184^{**}	0.001	-0.178	-0.112
									(0.031)	(0.123)	(0.166)	(0.075)
Duration of previous employment spell:												
$<$ 3 m.									-0.301^{**}	-0.408^{**}	0.132	-0.106
									(0.042)	(0.110)	(0.119)	(0.100)
$3-6$ m.									-0.323^{**}	-0.234^{*}	0.147	-0.173
									(0.042)	(0.098)	(0.085)	(0.094)
$>$ 1 yr									-0.313^{**}	-0.274^{**}	0.062	-0.176^{*}
									(0.035)	(0.088)	(0.062)	(0.071)

Notes: The model equations are given by (1) to (3), the likelihood is given by (6). SE in parentheses. * : $p < 0.05$ and ** : $p < 0.01$. Model 3 covariates (z) also include demographics (sex, married, number of children, and age group dummies). Reference category for employment durations: $3-6$ months. 'order of unemployment spell' refers to the second, third etc. unemployment spell.

4.1 The causal effects of becoming unemployed on return migration hazards

The estimated causal effects of unemployment spells on return migration hazards are reported in Table 4. Across all three specifications and all immigrant groups it is evident that unemployment dynamics shorten migration durations.

The average effect γ_u for the all groups of immigrants is substantial and essentially of the same magnitude, the point estimates ranging from 0.46 to 0.67. Models 2 and 3 reveal that the causal effect exhibits duration dependence. For EU migrants, the impact peaks for durations of 3-6 months.[6] For non-EU migrants, the picture is more heterogeneous, as duration dependence increases for immigrants from LDCs, whereas the coefficients become smaller for the others. Further permitting the causal effect to vary across characteristics (demographics and labour market history) increases the magnitude of the duration effects for EU immigrants. This follows, in particular, since the duration of the preceding employment spell lengthens the migration spell. By contrast, the effect of the previous labour market history is found to be insignificant for non-EU DC immigrants.

4.2 The causal effects of becoming re-employed on return migration hazards

Finding employment after a period of unemployment is a positive labour market event which is likely to impact also on migration durations. Table 5 reports the results. For all except new EU immigrants, the effect of having found employment after an unemployment spell delays the migrant's return. The effect is particularly strong for immigrants from developed countries outside the EU. Previous unemployment durations only exhibit an effect if these were no longer than 3 months, indicating that such unemployment spells were anomalies which were quickly overcome by the individual. The one immigrant group which deviates from this pattern of extended migration durations are immigrants from the new EU, i.e. predominantly Polish immigrants. The estimated causal impact of re-employment for this group, however, is consistent with target savings: having re-gained employment, it is plausible that such immigrants are back on track to reach their savings target and return once this has been attained.

[6]The point estimates in Model 3 for EU15 for 3-6 months, .963, and in excess of 1 year, 1.0, are statistically not distinguishable.

Table 5: Estimated causal effect of re-employment on return migration hazards.

	Model 1				Model 2				Model 3			
	EU 15	new EU	non-EU		EU 15	new EU	non-EU		EU 15	new EU	non-EU	
			DCs	LDCs			DCs	LDCs			DCs	LDCs
Constant effect [γ_e]	−0.118**	0.252**	−0.441**	−0.161*								
	(0.037)	(0.067)	(0.108)	(0.080)								
Duration dependence [α_e]:												
(0-3 months)					−0.095	0.281*	−0.271	−0.309*	−0.081	0.339	−0.178	−0.291
					(0.063)	(0.111)	(0.181)	(0.146)	(0.080)	(0.204)	(0.250)	(0.182)
(3-6 months)					0.109	0.441**	−0.122	0.041	0.143	0.536**	−0.041	0.098
					(0.063)	(0.119)	(0.182)	(0.144)	(0.081)	(0.208)	(0.251)	(0.181)
(6-12 months)					0.063	0.202	−0.105	−0.016	0.118	0.333	−0.042	0.072
					(0.055)	(0.116)	(0.143)	(0.127)	(0.075)	(0.209)	(0.225)	(0.168)
(> 1 year)					−0.295**	0.138	−0.598**	−0.288**	−0.197**	0.334	−0.554**	−0.106
					(0.046)	(0.107)	(0.115)	(0.108)	(0.070)	(0.209)	(0.208)	(0.154)
Labour market history [ϕ_e]:												
Repeated re-employment									0.040	0.218	−0.281	0.224
									(0.085)	(0.201)	(0.496)	(0.190)
Order of re-employment spell (> 1)									−0.139**	−0.025	−0.107	−0.041
									(0.045)	(0.114)	(0.336)	(0.095)
On benefit									−0.446**	−0.881**	0.108	−0.575**
									(0.084)	(0.313)	(0.271)	(0.210)
Duration of previous unemployment spell:												
< 1 m.									−0.382**	−0.323*	−0.085	−0.246
									(0.077)	(0.162)	(0.277)	(0.169)
$1-2$ m.									−0.279**	−0.139	0.221	−0.600**
									(0.077)	(0.169)	(0.241)	(0.191)
$2-3$ m.									−0.172*	0.041	−0.056	−0.225
									(0.077)	(0.167)	(0.242)	(0.170)
$6-12$ m.									0.032	−0.196	0.178	−0.301
									(0.078)	(0.186)	(0.219)	(0.183)
>1 yr									0.108	−0.156	−0.449	−0.161
									(0.098)	(0.232)	(0.295)	(0.235)

Notes: As for Table 4.

4.3 Impacts on immigration durations

We proceed to illustrate the impact of labour market histories on immigration durations. Specifically, we take the coefficient estimates of the return migration hazard models λ_m and consider, for each immigrant group, several labour market profiles. For simplicity, we abstract first from observable individual heterogeneity and focus on the 'reference' immigrant by setting the covariate coefficients β_x^m, ϕ_u, and ϕ_e in equation (3) to zero. The object of interest is the survival probability $Pr\{T_m > t\} = E_{v_m}\{\exp\left(-\int_0^t \lambda_m(s)\, ds\right)\}$ where the expectation is taken over unobserved heterogeneity v_m, recalling that T_m is the sojourn time in the host country. We consider variations in unemployment durations, in the timing of unemployment spells, and conclude with an analysis of counterfactuals.

4.3.1 The impact of unemployment durations

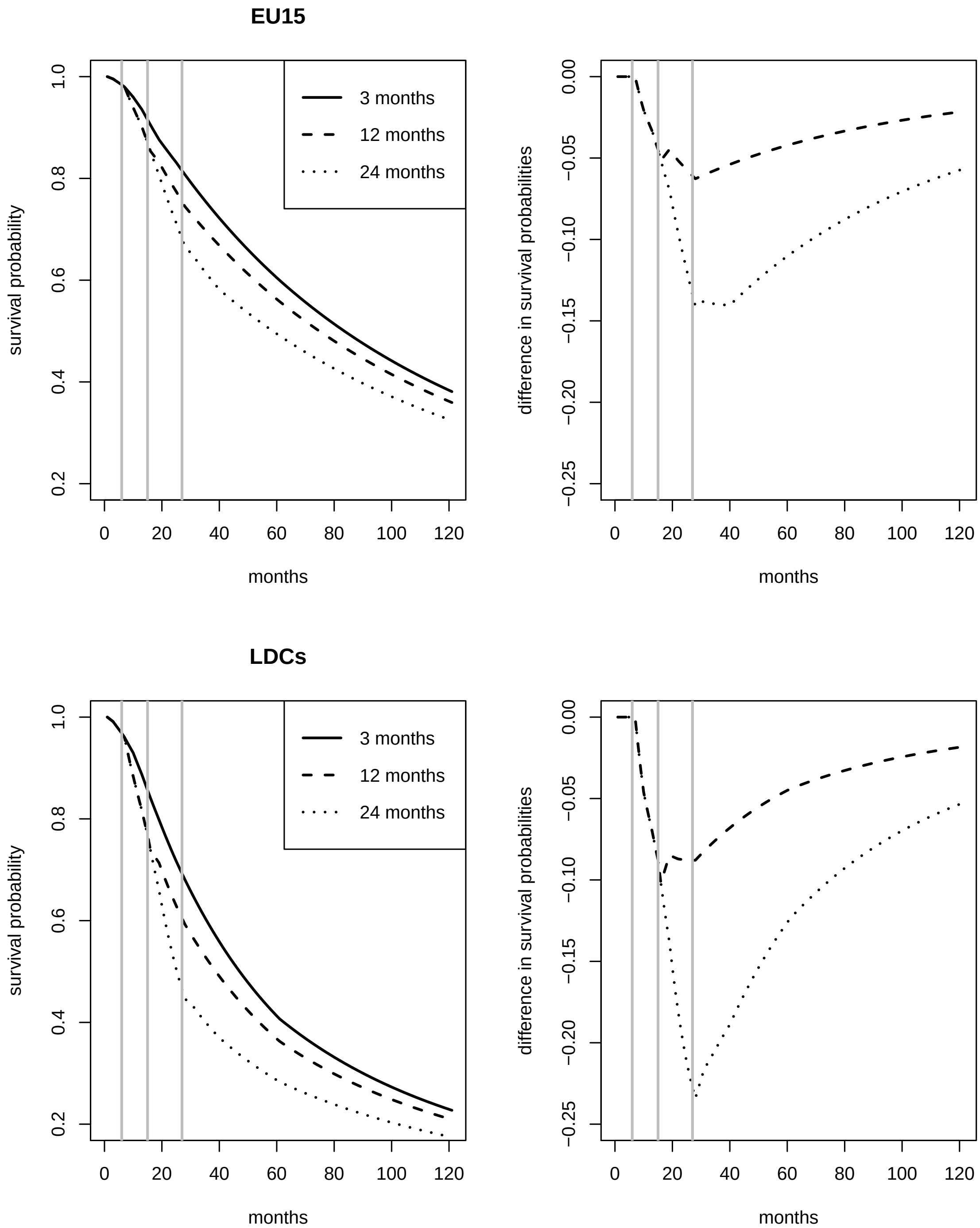
We compare the impact of unemployment durations of 3, 12, and 24 months, whereupon immigrants, who are still in the host country, experience a reversal of fortunes and find employment again. To ensure comparability, all unemployment spells start three months after entry into the Netherlands.

Figure 3 depicts the results for immigrants from the EU15 and from the non-European LDCs. The three vertical lines indicate the end of the respective unemployment spell, and thus also indicate the time interval where survival probabilities coincide since unemployment spells coincide. To facilitate the comparison between the survival probabilities, we also plot in the left panel the difference between the survival probabilities associated with the longer unemployment durations and with the duration of 3 months.

The figure illustrates again, first that unemployment spells shorten survival probabilities, and second that the magnitude of the impact increases in the unemployment duration. Relative to the survival probabilities associated with the 3 month unemployment spell, the largest difference occurs at the time the respective unemployment spell comes to an end (at times 3+12 and 3+24 months), whence survival probabilities start to converge again.

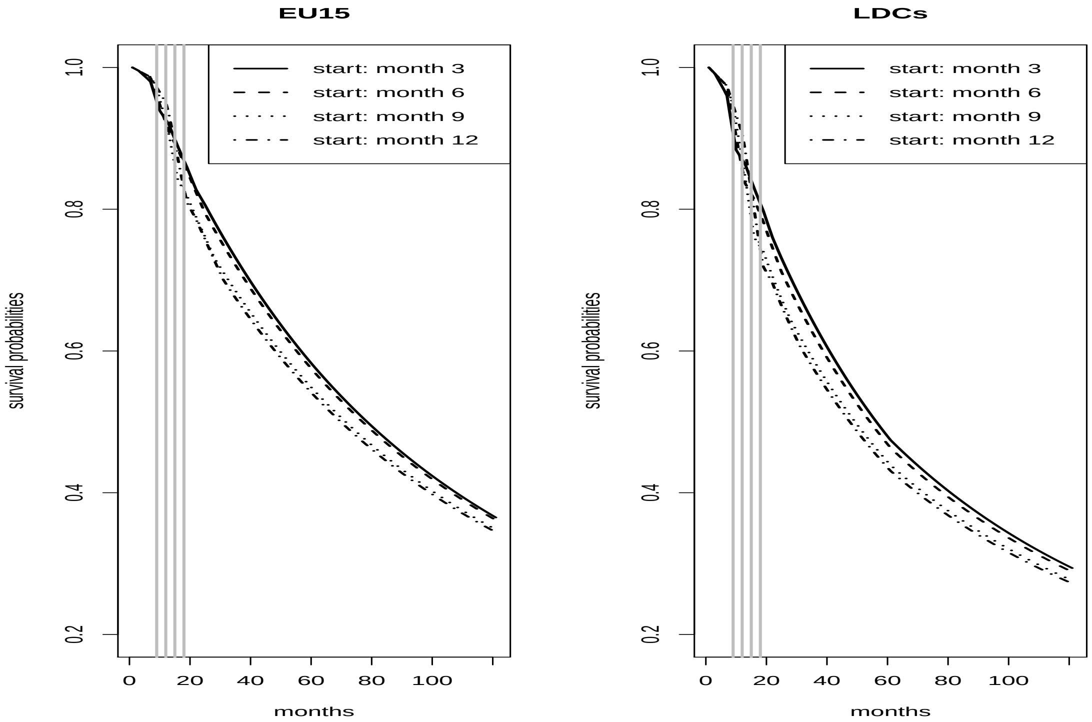
Finally, the figure also illustrates the difference between the different immigrant groups, both in terms of the absolute values of the survival probabilities (left panels), as well as the relative differences (right panels). In particular, the maximal difference in survival probabilities for immigrants from the EU15 (LDCs) for unemployment durations 3 and 12 is 0.06 (0.1), and for the unemployment durations of 3 and 24 months is 0.14 (0.23). Hence the impact of longer unemployment durations from immigrants from LDCs is substantially larger.

Figure 3: The effect of unemployment spells on survival probabilities



4.3.2 The effect of the timing of the unemployment spell

Figure 4: The effect of the timing of unemployment spells on survival probabilities



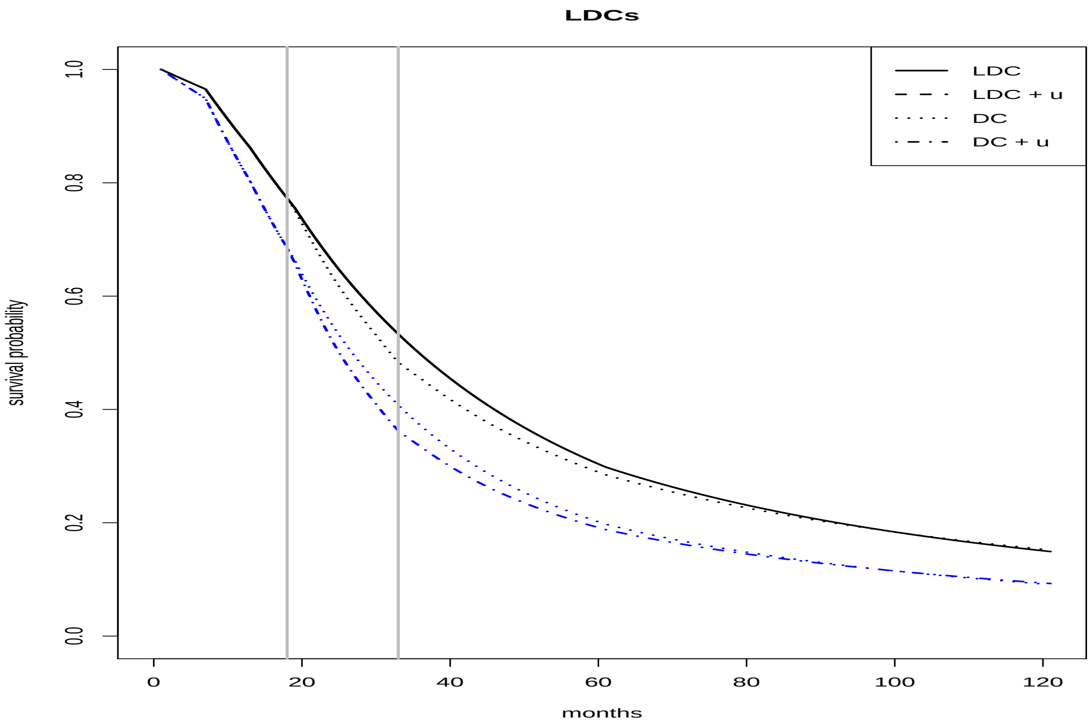
In order to assess the effect of the timing of an unemployment spell, we consider spells of a common length of 6 months, with start times at 3, 6, 9, and 12 months after entry into the Netherlands. Figure 4 displays the results. For survivals of at least 20 months, it is clear that while later starts of the unemployment spells have larger impacts on the survival probabilities, the differences are fairly small. For instance, for immigrants from the EU15, the maximal difference between the survival probabilities associated with the earliest and the latest start of the unemployment spell is .06, whereas this becomes .08 for immigrants from LDCs. Hence this difference between the immigrant groups is fairly small.

4.3.3 Counterfactual analysis

As policy makers in developed host countries often seek to attract higher 'quality' immigrants, it is of interest to consider the impact of imputed 'quality improvements' on survival probabilities by essentially comparing outcomes for immigrants from non-European LDCs with average LDC characteristics and with average DC characteristics (holding thus constant the covariate coefficients of the return migration hazard model). Compared to the two previous exercises, the coefficient vector β_x^m is not zero and given by the estimates for the model for immigrants from LDCs. Average covariate values are reported in Table 8 in the Data Appendix. In particular, we consider single male wage earners in the service sector who do not own a house, who are averaged aged (31.6 vs. 35), and who earn average wages when employed (€ 2,751 vs. € 5,476).

While covariates have a direct impact on survivals, they also, of course, affect unemployment propensities. To isolate the direct covariate effect, in Figure 5 the lines labelled 'LDC'

Figure 5: Counterfactual analysis: LDC immigrants with average DC characteristics



and 'DC' depict the survival probabilities for LDC immigrants with the respect covariate profile who are continuously employed. The additional unemployment effect is indicated by the line '+u' and computed as follows: Based on the results of Tables 1 and 2, we compute the expected survival probability, weighted by the probability of being always employed (62.1% vs. 64%). For the survival probability conditional on having been unemployed, we assume that the unemployment spell starts at the average first employment length (17 months), and lasts the average length (13 months). As Table 2 revealed, there is little difference between immigrants from LDCs and DCs in terms of the last two outcomes.

Figure 5 reveals that the unemployment effect is very small, and the difference in profiles is driven by the direct covariate effect, which is dominated by the difference in average earnings. If the considered average immigrant from LDCs were to earn, on average, as immigrants from DCs when employed, then immigration durations would be shorter, but the reduction in survival probability would not exceed 13 percentage points. It is in this sense that improved immigrant 'quality' would reduce immigration durations.

4.4 Selectivity biases

We briefly illustrate the consequences of ignoring the endogeneity issue induced by the correlations between the unobservable heterogeneity terms (v_e, v_u, v_m) given by distribution G. Recall that we have modelled this by equation (4). We quantify the resulting selectivity biases in terms of the principal objects of interest, namely the point estimates of the causal effects on return migration hazards. For the sake of brevity, we illustrate the issue in the context of Model 1, i.e. the estimate of the constant causal effect given by γ_k with $k \in \{u, e\}$.[7] Table 6 reports the results.

It is evident that ignoring the endogeneity results in positive biases of the causal effect of unemployment for all groups except the non-EU DC group; in the case of re-employment, the causal effects are also reduced in magnitude. In the former case, the biases range between 9% and 33 %, in the latter case the range is from 34% to 46% except for non-EU LDCs in which case the bias factor is close to 2.5. We conclude that not controlling for selectivity results in substantial biases.

Table 6: Analysis of selectivity biases on the effect on return-migration hazards

	EU 15	new EU	non-EU DCs	non-EU LDCs	EU 15	new EU	non-EU DCs	non-EU LDCs
	Unemployment				*Re-employment*			
Timing of Events model	.638**	.670**	.464*	.667**	−.118**	.252**	−.441**	−.161*
	(.076)	(.165)	(.185)	(.193)	(.037)	(.067)	(.108)	(.080)
PH-model	.771**	.787**	.318	.673**	−.179**	.177**	−.345**	−.397**
	(.065)	(.154)	(.163)	(.152)	(.026)	(.061)	(.076)	(.060)
MPH-model	.851**	.820**	.440*	.732**	−.162**	.187**	−.301**	−.384**
	(.069)	(.156)	(.173)	(.158)	(.028)	(.062)	(.080)	(.130)

Notes: $^{*}p < 0.05$ and $^{**}p < 0.01$. PH refers to the P(roportional)H(azard)-model. The M(ixed)P(roportional)H(azard)-model has a discrete unobserved heterogeneity distribution, and is given by (3) alone with $\alpha_u = \alpha_e = \phi_u = \phi_e = 0$ and thus ignores the correlation with the error terms appearing in equations (1) and (2) .

5 Conclusion

The majority of recent labour immigration to the Netherlands is temporary rather than permanent. Across all immigrant groups, a substantial proportion leave the host country eventually, and many do so within 24 months. We have considered in this paper the individual labour market drivers of immigration durations.

Despite this extent of temporary immigration, the interdependence of labour market events and immigration durations has received little attention in the empirical literature, mainly because of severe data limitations. We have addressed this gap using a unique Dutch administrative panel of the entire population of recent labour immigrants. Hence the usual concerns about immigrant data (small samples, missing covariates, latent migrant types, inaccurate measurement and recall) are absent, as we observe entry, exit, migration motive, and complete labour market histories. Moreover, the large size of the data enables us to estimate separate models for distinct immigrant groups, and we have shown the importance of controlling for observable migrant heterogeneity.

[7] For the sake of brevity, we do not report estimates of the factor loaders, and the implied covariances. These are, of course, available from the authors.

The principal methodological challenge arises, however, from unobservable heterogeneity that is correlated across the migration and the labour market processes. The timing of events method enables us to control for the selectivity of returnees, and thus to identify and estimate the causal effects of employment and unemployment histories on migration durations. Simpler models which ignore error correlations across labour market and migration processes are shown to exhibit substantial selection biases.

Overall, we have found that, across all immigrant groups, unemployment dynamics shorten the migration duration, while employment spells following unemployment spell delay the return for all migrants except for those from the new EU countries. The causal impact of labour market dynamics is quantified in terms of migration durations in several experiments, focussing on the duration and timing of unemployment spells, and, in a counterfactual analysis, the effect of improved immigrant "quality". These experiments show that the unemployment durations have a substantial effect, while the effect of differences in timing and "quality" are relatively smaller.

References

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

Bijwaard, G. E. (2009) "Labour Market Status and Migration Dynamics", IZA Discussion paper 4530, IZA Bonn.

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

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

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

Dustmann, C. (2003) "Children and return migration", *Journal of Population Economics*, 16: 815–830.

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

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

Gibson, J. and D. McKenzie (2009) "The Microeconomic Determinants of Emigration and Return Migration of The Best and Brightest", Research Working papers.

Kirdar, M. (2007) "Labour market outcomes, capital accumulation, and return migration: Evidence from immigrants in Germany", MPRA Paper 2028, University Library of Munich, Germany.

Yang, D. (2004) "Why Do Migrants Return to Poor Countries? Evidence from Philippine Migrants' Responses to Exchange Rate Shocks", Working Papers 513, Research Seminar in International Economics, University of Michigan.

A Data Appendix

Table 7 explains the sub-populations of labour immigrants, and disaggregates these according to the country of birth. For countries outside the EU, we distinguish between developed (DCs) and less developed countries (LDCs). Among the EU15, immigrants from the UK and Germany predominate, for the new EU 69% are Polish immigrants. In the non-EU group, the three largest subgroups are from the USA, India and Japan. Overall immigrants from non-EU LDCs are the second largest group.

This grouping also corresponds to the varying degrees of labour mobility among the four groups. Immigrants from the EU15 can move freely in the Dutch labour market, as can, since 2004, immigrants from the new EU except for Bulgarians and Rumanians. All non-EU migrants need a work permit (the "Machtiging Voorlopig Verblijf (MVV)" or "Regular Provisional Residence Permit"). LDCs and DCs differ in that immigrants from these DCs are exempted from obtaining this MVV before entry. To obtain a work permit, three conditions must be met: (i) the presence of prioritised supply (i.e. a labour market check), and the recruitment efforts of the employer to fill the position with a native; (ii) renumeration in accordance with the market, and at least at the level of the statutory minimum wage; (iii) having secured adequate accommodation. Although self-employed migrants are exempted from the work permit requirement, residence permits are only granted if the authorities deem that the immigrant would serve 'vital' Dutch interest. The passage of time increase labour mobility of non-EU immigrants. In particular, the migrant will be entirely free to work on the Dutch labour market after 3 years of residence on the basis of a residence permit to perform work. After 5 years of residence the migrant can apply for a permanent residence permit (or citizenship).

Table 7: Major country of birth

EU 15		new EU		non-EU			
					DCs	LDCs	
UK	27.4%	Poland	69.1%	USA	38.5%	India	19.2%
Germany	18.5%	Romania	10.1%	Japan	26.8%	China	10.2%
France	9.3%	Czechoslovakia	7.4%	Australia	10.1%	South Africa	7.8%
Portugal	8.4%	Hungary	6.0%	Canada	8.1%	Brasil	3.7%
Italy	8.2%	Bulgaria	5.4%	South Korea	4.9%	Taiwan	3.5%
Belgium	7.1%	Lithuania	0.8%	Norway	4.1%	Morocco	2.9%
Spain	5.7%			Switzerland	3.8%	rest Africa	17.6%
Greece	4.3%			New Zealand	3.1%	rest Asia	13.9%
Ireland	3.0%					Latin America	10.0%
Sweden	2.9%						
Denmark	1.8%						
Finland	1.8%						
Austria	1.3%						
$N =$	48,290	12,717		11,746		16,974	

Table 8 reports some summary statistics for the sub-populations. The migrant groups look fairly similar in terms of the age distribution and occupational choices. Relative to EU migrants, non-EU DC migrants are more often male, married, less often own a house, and have a substantially larger share among the highest income group (recall that we require that the labour immigrant be employed at first entry). Remarkable about migrants from the new EU states (69% of whom are Polish) is the fact that earnings are fairly homogeneous, with 36% in the first and 44% in the second income group.

Table 8: Descriptive statistics at first entry

	EU15	new EU	non- EU DCs	non- EU LDCs
# of migrants	48,290	12,717	11,746	16,974
Female	33.1%	29.6%	22.2%	23.3%
Single	80.3%	72.7%	56.8%	71.3%
Married	17.1%	25.9%	42.2%	27.5%
Divorced	2.5%	1.3%	0.9%	1.2%
Other origin	16.2%	15.4%	19.6%	37.0%
Age				
18–25	20.5%	26.1%	9.5%	16.0%
25–30	29.7%	31.0%	22.7%	31.9%
30–35	21.1%	18.0%	22.4%	23.0%
35–40	12.4%	10.0%	17.3%	13.1%
40–45	7.6%	6.9%	12.0%	8.2%
45–50	4.5%	5.0%	7.8%	4.6%
50–55	2.7%	2.2%	5.3%	2.2%
55–60	1.2%	0.7%	2.5%	0.8%
60–65	0.2%	0.1%	0.5%	0.2%
Average	31.3	30.3	35.0	31.6
Housing				
house owned	34.1%	33.9%	27.4%	28.1%
value house $<$ 100k	30.6%	19.5%	14.9%	26.7%
value house 100–200k	38.0%	36.3%	37.8%	37.5%
value house 200–300k	13.0%	16.1%	19.8%	15.6%
value house 300–400k	5.2%	6.0%	10.7%	5.4%
value house $>$ 400k	5.3%	8.1%	10.6%	6.4%

Table 8: Descriptive statistics at first entry (continued)

	EU15	new EU	non-EU DCs	non-EU LDCs
monthly income				
< 0	0.2%	0.5%	0.2%	0.1%
0 - 1000	32.5%	35.9%	21.5%	31.1%
1000 - 2000	26.2%	44.2%	12.0%	24.0%
2000 - 3000	18.2%	12.3%	12.7%	16.8%
3000 - 4000	7.6%	3.7%	8.9%	9.9%
4000 - 5000	3.9%	1.3%	6.8%	5.2%
5000 - 6000	2.9%	0.7%	6.0%	3.2%
> 6000	8.5%	1.4%	31.9%	9.7%
Average	€2517	€1484	€5476	€2751
Sector				
Agriculture	1.1%	5.7%	0.4%	1.2%
Industry	11.9%	9.6%	12.5%	8.9%
Construction	1.8%	1.9%	0.5%	0.6%
Catering	6.3%	2.1%	3.8%	4.2%
Trade	13.4%	9.9%	20.1%	10.8%
Transport	5.2%	3.5%	7.1%	3.8%
Finance	3.0%	1.8%	5.9%	3.7%
Services	42.5%	43.7%	33.9%	49.5%
Education	6.0%	5.5%	6.7%	10.1%
Care	2.8%	1.8%	2.0%	3.0%
Nonprofit	2.9%	1.7%	5.5%	2.6%