

**NORFACE**  
MIGRATION



---

**NORFACE MIGRATION** Discussion Paper No. 2011-27

# International Migration with Heterogeneous Agents: Theory and Evidence for Germany, 1967-2009

Herbert Brücker and Philipp J.H. Schröder

# International Migration with Heterogeneous Agents: Theory and Evidence for Germany, 1967-2009

Herbert Brücker\*      Philipp J.H. Schröder†

January 2011

## Abstract

Temporary migration, though empirically relevant, is often ignored in formal models. This paper proposes a migration model with heterogeneous agents and persistent cross country income differentials that features temporary migration. In equilibrium there exists a positive relation between the stock of migrants and the income differential, while the net migration flow becomes zero. Consequently, existing empirical migration models, estimating net migration flows, instead of stocks, may be misspecified. This suspicion appears to be confirmed by our investigation of the cointegration relationships of German migration stocks and flows since 1967. We find that (i) panel-unit root tests reject the hypothesis that migration flows and the explanatory variables are integrated of the same order, while migration stocks and the explanatory variables are all  $I(1)$  variables, and (ii) the hypothesis of cointegration cannot be rejected for the stock model.

*Keywords:* International migration, temporary migration, panel cointegration. *JEL code:* C23, C53, F22.

---

\*University of Bamberg, IAB, Nürnberg, and IZA, Bonn.

†Aarhus School of Business and DIW Berlin.

Corresponding author: Herbert Brücker, IAB, Regensburger Str. 104, D-90478 Nürnberg, GERMANY. Fon: +49-911-179-3807, Email: herbert.bruecker@iab.de.

Acknowledgements: Herbert Brücker gratefully acknowledges financial support from the NORFACE project "Temporary Migration, Integration and the Role of Policies" (TEMPO).

## 1 Introduction

Today's international migration – opposed to that of the 19th century – consists predominantly of temporary migrants, where an individual's stay abroad may vary from a few months to several decades. Return migration in Europe makes up almost 10 per cent of migration stocks p.a.,<sup>1</sup> and micro studies indicate that around 80 per cent of the migrants in Europe eventually return to their home countries (Dustmann, 1995; Karras and Chiswick, 1999). A central implication of migration being temporary is that each migration stream is associated with a compensating counter-stream, as Ravenstein (1889) already observed in his famous 'Laws of Migration' at the end of the 19<sup>th</sup> century.

Although often quoted, many of Ravenstein's fundamental observations have not been addressed systematically in the migration literature. Most models in the classical migration literature treat migration as permanent. Starting with the seminal contributions of Hill (1987) and Djajic and Milbourne (1988) a number of models have analysed the phenomenon of temporary migration, but these models usually ignore the heterogeneity of individuals. Consequently, even though migration is temporary, the same length of migration episodes applies for all agents.<sup>2</sup> There exist however striking micro evidence that the length of migration episodes differs largely across individuals depending on their human capital characteristics (Dustmann, 1995,

---

<sup>1</sup>See the evidence provided by OECD (2010) for a number of developed countries; for Germany see Statistisches Bundesamt (2010b).

<sup>2</sup>Hill (1987) and Djajic and Milbourne (1988) treat migration as an intertemporal optimisation problem, where the length of migration is endogenously determined by host and home wages and differences in utility between consumption abroad and at home. Yet as they employ the concept of a representative agent all migration decisions (including the length of stay) are identical for all agents. Building on these models, Dustmann and Kirchkamp (2002) and Mesnard (2004) consider problems such as liquidity constraints, differences in purchasing power parities across countries, and enhanced options for self-employment for return migrants.

2003; Dustmann and Kirchkamp, 2002; Mesnard, 2004), although this evidence is not considered by the underlying migration models. One exception in the literature is the model of Stark (1995), which explains differences in migration duration by asymmetric information concerning the human capital characteristics of high and low productivity migrants. After the true type is revealed, low-productivity migrants are dismissed and return home, accordingly displaying a shorter migration duration than high productivity types.

This paper takes another route by considering heterogeneous preferences of individuals with regard to the choice of location. The basic set-up of the model is related to the standard model of temporary migration originally developed by Djajic and Milbourne (1988), but departs in several aspects: First, and most importantly, it is assumed that individuals discount consumption in foreign countries by a certain factor, which varies across individuals.<sup>3</sup> Even though not dealt with explicitly in the present paper, such discount will be determined by the migration drivers and barriers identified in the literature apart from income differences. For example it may increase in the presence of fixed migration costs (Grogger and Hanson, 2011), or fall when an existing migrant population abroad creates beneficial network effects (Beine, Docquier, and Çaglar Özden, 2011). Second, in order to arrive at analytical solutions for the case of heterogeneous agents with individual and endogenously determined length of migration spells, we have to employ specific functional forms. Finally, the present model departs from Djajic and Milbourne (1988) with respect to wage reactions to migration flows. Arguably, the traditional assumption where wage equalisation determines the

---

<sup>3</sup>Some other models in the literature consider heterogeneous preferences as well (Faini and Venturini, 1995), but to the best of our knowledge the consequences of heterogeneous preferences have not yet been analysed in the context of temporary migration.

amount of migration in equilibrium appears to run counter to empirical evidence. Actual migration is small relative to persistent cross-country income differentials. Furthermore, institutional settings in Europe do limit wage reactions in both the receiving and sending countries. Nevertheless migration stocks are relatively stable over time despite these large wage arbitrage opportunities, this is for example the case within the EU and the European Economic Area (EEA), where migration barriers have been largely abolished. Based on these observations, equilibrium in the present model is not driven by the wage equating forces of migration, but instead by the heterogeneity of agents, which determines both the amount and duration of migration in reaction to a given income differential.<sup>4</sup>

More specifically, at a given income differential the model generates three types of individuals: stayers, i.e. those who stay at home and do not migrate; temporary migrants, i.e. those who return home within their lifetime; and permanent migrants, i.e. those who migrate for their entire lifetime. Furthermore, within the group of temporary migrants, the duration of the migration episode varies across individuals. Accordingly migration flows and stocks are composed of a variety of different behaviors by agents. The average duration of migration episodes as well as the number of permanent migrants tend to increase with an increasing income differential. The stock of migrants, i.e. the share of the population which tends to stay abroad at a certain point of time, increases with the income differential between the host and the home country as well. Moreover, the net migration flow is zero in equilibrium. Gross migration flows remain, however, a positive function of the income differential.

These results have important implications for the estimation of macro

---

<sup>4</sup>Alternatively, the present analysis may be viewed as a partial equilibrium version of more traditional set-ups.

migration models: The standard macro models of migration, following the famous Harris and Todaro (1970) model, presume that an equilibrium between net or gross migration *flows* and the explanatory variables emerges. This tradition has inspired until today the specification of macro migration models, which explain migration flows by wage differentials, employment rates and a number of institutional variables, see e.g. the seminal contributions by Hatton (1995), Mayda (2010) and Pederson, Pytlikova, and Smith (2008). In contrast our model implies that an equilibrium relationship between migration *stocks* and the explanatory variables arises in the long run, while net flows converge to zero even in case of large and persisting income differences between countries. Accordingly, we test the hypothesis whether migration stocks or flows and the explanatory variables form a long-run equilibrium in the empirical part of the paper, or, in technical terms, a cointegration relationship.

Our analysis is based on migration to Germany from 51 source countries during the period 1967-2009. Germany is not only after the US and Russia the largest destination of migrants in the world in absolute terms, it is also one of a few countries which provides time series on migration stocks and flows for more than four decades. Following the Engle and Granger (1987) procedure, we test first whether the variables of the stock or the flow model form a cointegrated set. To this end we apply panel unit-root and panel cointegration tests, which increase the statistical power in comparison to univariate unit-root and cointegration tests. We find for our dataset at hand that migration flows are stationary variables while the explanatory variables such as income and employment variables are integrated of the first order (I(1) variables). Thus the hypothesis of the traditional migration model in the empirical literature that migration flows and the explanatory variables

form a cointegrated set is not supported by our data set. In contrast, we provide evidence that migration stocks are  $I(1)$  as well. Moreover, our panel cointegration tests rejects the null hypothesis of no cointegration for most specifications of our error correction model.

The remainder of this paper is as follows. Section 2 presents the formal model of migration with heterogeneous agents. It derives the amount of permanent and temporary migration as well as the individual and aggregate duration of migration episodes and presents results for the implied migration stocks and flows. Section 3 discusses the alternatives of flow and stock specifications for the estimation of macro migration models and applies then panel unit-root and panel cointegration tests in order to prove whether the variables of these alternative specifications form a cointegrated set. On this basis, Section 4 estimates the cointegrating vectors and the short-run dynamics of a migration stock model by employing an error correction model. Section 5 concludes.

## **2 A migration model with heterogeneous preferences**

### **2.1 The model**

Consider an economy where at each instant in time,  $t$ , there are  $N$  individuals  $i$  born, endowed with one unit of labour each, and who each live for the same period of time,  $T_i$ , normalised to 1, i.e.  $T_i = 1, \forall i = 1, \dots, N$ . Each individual is continuously employed throughout his or her life but has the choice of staying abroad for a period  $\tau_i$ , where  $0 \leq \tau_i \leq 1 \forall i = 1, \dots, N$ . As in Djajic and Milbourne (1988), agents make and execute their migration

decision at time  $t_i = 0$ . In the foreign country, each domestic and migrant worker receives the income level  $y^*$ , in the sending country the income level is  $y$ , where  $y < y^*$ , i.e. outward migration only occurs from home to foreign country.<sup>5</sup> The utility flows which individuals perceive from consumption (that is, living) at home and abroad respectively are given by:

$$u(c_i) = c_i^\alpha \quad (1)$$

$$u^*(c_i^*) = \gamma_i^{1-\alpha} c_i^{*\alpha}, \quad (2)$$

where  $c_i$  and  $c_i^*$  are consumption at home and abroad, respectively,  $\alpha$  ( $0 < \alpha < 1$ ) is a parameter of the utility function, identical for all agents, and  $\gamma_i \in [0, 1]$  is a preference parameter, which is heterogenous across agents.

The parameter  $\gamma_i$  captures the fact that individuals receive less utility from consumption abroad than at home. The utility discount embedded in  $\gamma_i$ , may represent real or perceived cost of living abroad, include the flow equivalent of fixed migration costs that migrant face in their migration decision (Beine, Docquier, and Çaglar Özden, 2011), and may vary with the presence of other natives living abroad, i.e. network effects (Grogger and Hanson, 2011). Although we do not explore these issues in the present model, in passing we note that different distribution functions of gamma, and changes of these distributions across time and destinations would be able capture these and potential other drivers and determinants of migration decisions.

The utility functions in (1) and (2) display the feature that the marginal utility enjoyed from the same rate of consumption is higher at home than abroad, i.e. that  $u^{*'}(x) < u'(x)$ , thus fulfill the conditions laid out in Djajic and Milbourne (1988).<sup>6</sup>

<sup>5</sup>Variables with an asterisk denote throughout the Section values in the foreign country.

<sup>6</sup>As is usual in the literature, we interpret the condition that  $u^{*'}(x) < u'(x)$  to capture

The lifetime utility of a migrating individual returning to the home country at time  $\tau_i$  can then be written as<sup>7</sup>

$$V_i = \tau_i \gamma_i^{1-\alpha} c_i^*(t)^\alpha + (1 - \tau_i) c_i(t)^\alpha. \quad (3)$$

The intertemporal maximisation problem of the individual is then straightforward (Djajic and Milbourne, 1988): choose the duration of the stay in the foreign country,  $\tau_i$ , and the rates of consumption over time abroad,  $c_i^*(t)$ , and at home,  $c_i(t)$ , such that lifetime utility (3) is maximised subject to the budget constraint

$$\tau_i y^* + (1 - \tau_i) y - \tau_i c_i^*(t) - (1 - \tau_i) c_i(t) \geq 0. \quad (4)$$

The first-order conditions (see Appendix A.1) give rise to the following relations:

$$\alpha \gamma_i^{1-\alpha} c_i^*(t)^{-(1-\alpha)} = \lambda, \quad (5)$$

$$\alpha c_i(t)^{-(1-\alpha)} = \lambda, \quad (6)$$

$$\gamma_i^{1-\alpha} c_i^*(t)^\alpha + c_i(t)^\alpha = \lambda (c_i^*(t) - c_i(t) + y - y^*), \quad (7)$$

$$\tau_i (y^* - y + c_i(t) - c_i^*(t)) = c_i(t) - y. \quad (8)$$

Since the shadow value of wealth,  $\lambda$ , is time-invariant, (5) and (6) imply that  $c_i^*(t) = c_i^*$  and  $c_i(t) = c_i$ ,  $\forall i = 1, \dots, N$ , i.e. consumption at home and abroad is constant over time. Moreover, equating the left-hand side of (5) and (6)

---

the fact that closer social relations to friends and relatives in the home country, a familiar cultural environment and other factors associated with the home country result in a higher utility for the same rate of consumption in the home country (Faini and Venturini, 1995).

<sup>7</sup>Notice that we ignore discounting by setting the discount factor implicitly to one. Nevertheless, none of the results below depend on this assumption, see e.g. Dustmann (2003), Dustmann and Kirchkamp (2002), and Mesnard (2004) for a similar approach. Discounting is, however, included in Djajic and Milbourne (1988).

gives:

$$c_i^* = \gamma_i c_i. \quad (9)$$

Thus consumption during the migrants stay abroad is a fraction of the consumption upon his/her return to the home country. Next, from (9) and (7), after substituting  $\lambda$  from (6), we are able to solve for  $c_i$ :

$$c_i = \frac{\alpha}{1 - \alpha} \frac{1}{1 - \gamma_i} (y^* - y), \quad (10)$$

i.e. consumption at home – and thus via (9) also consumption abroad – is a linear function of the income differential.<sup>8</sup>

Finally, using (9), (10) and (8) one can calculate the optimal length of a migrants's stay abroad:

$$\tau_i = \frac{\alpha}{1 - \gamma_i} - \frac{(1 - \alpha)y}{y^* - y}. \quad (11)$$

This optimal duration of migration displays the following reactions to changes in the various parameters (see Appendix A.2). With respect to the income levels, we find that  $\frac{\partial \tau_i}{\partial y} < 0$  and  $\frac{\partial \tau_i}{\partial y^*} > 0$ , thus an increase in foreign income, a reduction in the domestic income and hence a widening of the income gap leads to longer migration periods for all migrants. Furthermore, as one would expect intuitively,  $\frac{\partial \tau_i}{\partial \gamma_i} > 0$ , namely, individuals who have less of a utility discount when consuming abroad display longer migration duration. And the same follows for migration drivers and barriers that affect  $\gamma_i$ , for example will positive migrant network effects increase  $\gamma_i$  and thus migration spells.

Equations (9), (10) and (11) characterise the agents' migration and con-

---

<sup>8</sup>Notice that the consumption patterns established in (10) and (9) also define the savings path, e.g.  $s_i^*(t) = y^* - c_i^*$  for  $t = 0, \dots, \tau_i$ .

sumption behavior in the economy and are largely in line with results found in the literature following Hill (1987) and Djajic and Milbourne (1988).<sup>9</sup> On this basis, we are now equipped to explore the consequences of agent heterogeneity for aggregated migration patterns.

## 2.2 Results

Given the above model we can derive results concerning agent heterogeneity, migration decisions and migration flows and stocks.

### Permanent and temporary migrants and stayers

The optimal value of  $\tau_i$  given in (11) may well be larger than an agent's total lifetime,  $T_i = 1$ . This becomes more likely for very high  $\gamma_i$ , an individual with a small utility discount when living abroad, or for  $\alpha$  close to 1 or for a very large income gap  $y^* - y$ . In fact, what a  $\tau_i \geq 1$  implies is that an agent becomes a permanent migrant: the utility value of living and consuming abroad is so large that given the higher income level in the foreign country, returning – even in the last instant of life – creates no additional value. Define by  $\bar{\gamma}$  the individual who is indifferent to the question of returning (temporary migration) vs. staying abroad forever (permanent migration). Solving  $\tau_i = 1$  from (11) for  $\gamma$  gives the first result:

**Lemma 1.** *The group of permanent migrants consists of all individuals  $i$  with*

$$\gamma_i \geq \bar{\gamma} = \frac{(1 - \alpha)y^*}{y^* - \alpha y}. \quad (12)$$

---

<sup>9</sup>One important difference does exist, however: Djajic and Milbourne (1988) find an ambiguous effect of foreign income on the migration duration, since a higher income might result in an earlier return to the home country if the utility function is characterised by a low rate of substitution. This case is excluded here through the specific functional form of utility.

Solving for the consumption volume of a permanent migrant from (9) and (10) after setting  $\gamma_i = \bar{\gamma}$ , one can verify that  $c_i^* = y^*$ , i.e. permanent migrants spend their total income in the foreign country and do not save. Furthermore,  $\frac{\partial \bar{\gamma}}{\partial y^*} < 0$ , such that an increase in the foreign income level – or an increase in the income gap – lowers the threshold value of  $\gamma_i$ , beyond which individuals become permanent migrants.

At the other end of the spectrum, we have those agents who prefer to stay at home instead of migrating. Define by  $\underline{\gamma}$  the individual who is completely indifferent to the question of migrating vs. staying at home, i.e. the individual whose optimal migration duration is  $\tau_i = 0$ . Solving  $\tau_i = 0$  from (11) for  $\gamma$  gives:

**Lemma 2.** *The group of stayers (non-migrants) consists of all individuals  $i$  with*

$$\gamma_i \leq \underline{\gamma} = \frac{y - \alpha y^*}{(1 - \alpha)y}. \quad (13)$$

Since  $\gamma_i \in [0, 1]$ , a necessary condition for at least one individual in the sense of lemma 2 to exist is that  $y^* < y/\alpha$ . If this condition is violated, then the income gap is so substantial, that all individuals of the sending country would migrate. For the sake of realism it is assumed that  $y^* < y/\alpha$  is fulfilled in the remainder of the Section. Notice that  $\frac{\partial \underline{\gamma}}{\partial y^*} < 0$ , such that an increase in the foreign income level lowers the threshold value  $\underline{\gamma}$ , implying that fewer agents are stayers.

It is easy to verify that  $\bar{\gamma}$  in (12) is always larger than  $\underline{\gamma}$  in (13) as long as the income gap  $y^* - y$  is positive. Accordingly, there exists a third group of agents that maximise utility with a  $\tau_i \in ]0, 1[$ , i.e. individuals who spend part of their working lives abroad and part at home – temporary migrants. Following the reasoning above, temporary migrants are characterised as follows:

**Lemma 3.** *The group of temporary migrants consists of all individuals  $i$  with*

$$\gamma_i \in ]\underline{\gamma}, \bar{\gamma}[.$$

Thus, the above results establish that within the population of the home country, three types of agents can be distinguished. While all permanent migrants display identical consumption and migration durations – as do stayers – the group of temporary migrants features varying durations of migration spells. Figure 1 plots the qualities implied by the optimal migration duration from (11) and lemma 1, 2 and 3.

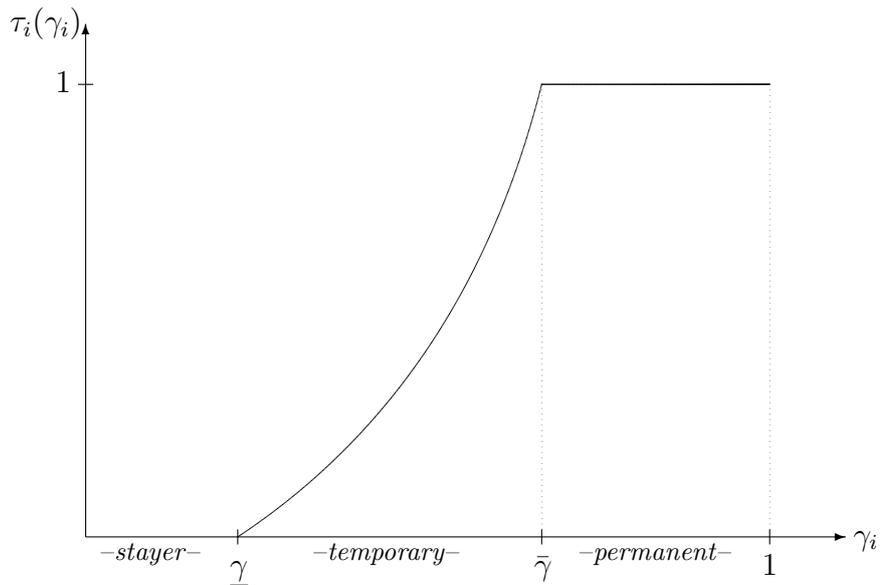


Figure 1: The duration of migration spells for different agents

### The aggregate duration of migration

In order to analyse aggregate effects, the distribution of  $\gamma_i$ 's in the population must be specified. Here we assume the  $\gamma_i$ 's in each cohort to be uniformly

distributed on support  $[0,1]$ . Accordingly the area under the curve in Figure 1 represents the aggregate duration of all migration spells for a cohort. While  $\tau_i$  is by definition zero for stayers and one for permanent migrants, Figure 1 shows that the migration duration is monotonically increasing in  $\gamma$  in the interval between  $\bar{\gamma}$  and  $\underline{\gamma}$ .

Integrating (11) over the interval  $\underline{\gamma}$  to  $\bar{\gamma}$  gives the aggregate duration  $\theta^t = \int_{\underline{\gamma}}^{\bar{\gamma}} \tau_i$  of temporary migrants:

$$\theta^t = \alpha N \left( \ln \left( \frac{y^* - y \alpha}{y (1 - \alpha)} \right) - \frac{y^* - y}{y^* - \alpha y} \right). \quad (14)$$

The resulting  $\theta^t$  captures the total migration duration of all temporary migrants of a cohort, thus in effect both the number of temporary migrants as well as their individual migration durations are captured. As one would intuitively expect, it can be shown that  $\frac{\partial \theta^t}{\partial y^*} > 0$ , i.e. an increase in foreign income levels has an unambiguously positive effect on the duration of aggregated temporary migration.

Next, under the assumption of uniformly distributed  $\gamma_i$ 's, the duration of aggregated permanent migration of a cohort simply becomes

$$\theta^p = 1(1 - \bar{\gamma})N = \frac{(y^* - y)\alpha N}{y^* - \alpha y}. \quad (15)$$

Finally, the number of stayers – and since  $T_i = 1$  also their aggregate time spent at home – found in a cohort is simply  $\underline{\gamma}N = \frac{(y - \alpha y^*)N}{(1 - \alpha)y}$ .

Combining (14) and (15), we obtain the following result:

**Proposition 1.** *The aggregate duration of migration,  $\theta$ , from a single cohort is*

$$\theta = \alpha N \ln \left( \frac{y^* - \alpha y}{y (1 - \alpha)} \right), \quad (16)$$

- increases in the foreign income level,  $\frac{\partial \theta}{\partial y^*} > 0$ , and
- falls in the domestic income level,  $\frac{\partial \theta}{\partial y} < 0$

Equation (16) arrives at a surprisingly simple specification of the total duration of time spent abroad by the migrants in a given cohort. Lemma 1, 2 and 3 and proposition 1 have clear implications for migration stocks, migration flows and their interaction.

### Migration flows and migration stocks

Moving from the migration decisions and durations in a single cohort to migration stocks, one has to specify the number of cohorts coexisting at any instant in time. Let  $L$  denote this number. Assuming zero population growth – that is a rate of reproduction of 1 – then the total population at any point in time is  $LN$ . Furthermore, assume that each cohort is identical to the previous including their consumption and migration decisions but that descendants'  $\gamma_i$ 's are uncorrelated to their parents  $\gamma_i$ 's and that reproduction takes place at the end of an agents lifetime.<sup>10</sup> This leads to the following results:

**Proposition 2.** *The population stocks at every instant in time are*

$$a) \text{ stock of permanent migrants: } S^p = LN(1 - \bar{\gamma}) = \frac{LN\alpha(y^* - y)}{y^* - \alpha y}.$$

$$b) \text{ stock of temporary migrants: } S^t = \frac{\theta^t}{N(\bar{\gamma} - \underline{\gamma})} LN(\bar{\gamma} - \underline{\gamma}) \\ = \frac{-\alpha NL}{y^* - y \alpha} \left( y^* - y - (y^* - y \alpha) \ln \left( \frac{y^* - y \alpha}{y(1 - \alpha)} \right) \right).$$

---

<sup>10</sup>This last assumption implies that temporary migrants give birth after they returned to the home country, while only permanent migrants give birth abroad.

c) total stock of migrants:  $S = S^p + S^t = \alpha NL \ln \left( \frac{y^* - \alpha y}{y(1-\alpha)} \right)$ .

and

d) total stock of home population:  $H = NL - S(t)$   
 $= NL \left( 1 - \alpha \ln \left( \frac{y^* - \alpha y}{y(1-\alpha)} \right) \right)$ .

*Proof (sketch):* Proposition 2 a) is the aggregate of all permanent migrants in one cohort times the number of cohorts coexisting at every point in time. Since we have normalized the agent's lifetime to one, this turns out to be  $S^p = L\theta^p$ . Proposition 2 b) is the average duration of the migration spell of a temporary migrant,  $\frac{\theta^t}{N(\bar{\gamma} - \underline{\gamma})}$  times the total number of temporary migrants coexisting at every instant in time,  $LN(\bar{\gamma} - \underline{\gamma})$ . It follows that  $S^t(t) = L\theta^t$ . Finally, proposition 2 c) follows from a) and b) and thus  $S = L\theta$ , and proposition 2 d) follows from c). Thus the stock of migrants is here a fairly simple logarithmic relation of the income gap.

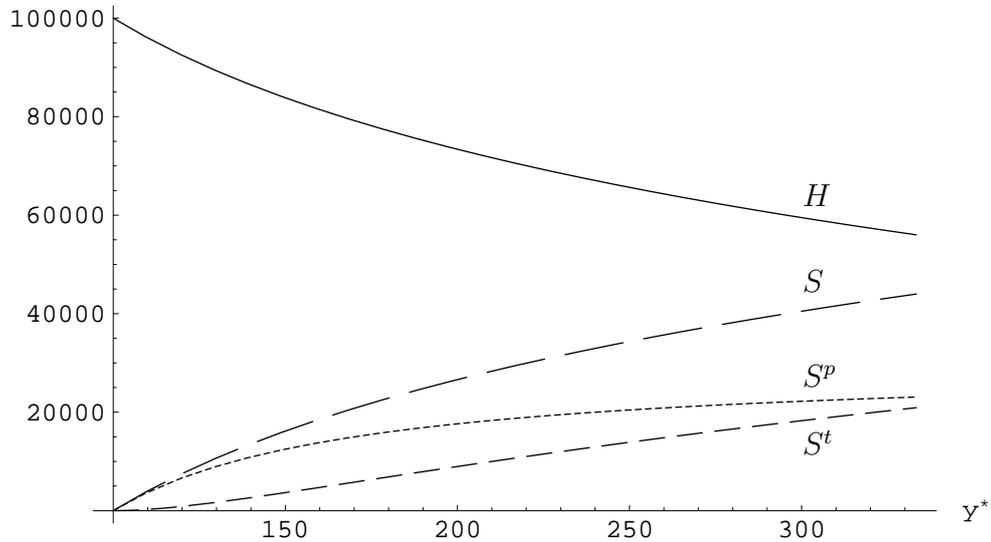


Figure 2: Migration stocks and stock of home population

To see what the relations derived in proposition 2 imply consider Figure 2. Figure 2 plots the stock of temporary, permanent, and total migration as well as the stock of home population as a function of foreign income,  $y^*$ . For the actual plot, the following parameter values are employed:  $y = 100$ ,  $\alpha = 0.3$ ,  $N = 1000$  and  $L = 100$ . At  $y^* = 100$  there is no income gap and accordingly all agents spend all their working life at home. As the income gap widens, there are initially a few individuals who opt for permanent migration and a few who opt for temporary migration. However, since the income differential is small, the actual amount of time spent abroad is small too. Accordingly from the perspective of the aggregate migration stock, temporary migration contributes relatively little to total migration compared to permanent migration since permanent migrants spend their entire lifetimes abroad. As the income gap widens, the role of temporary migration increases while that of permanent migration decreases.

Finally, consider the migration flows associated with the above stocks, in particular measuring flows occurring during any time interval of length 1. Given that reproduction takes place at the end of an agents life, the number of birth abroad occurring over the time interval 1 are  $LN(1 - \bar{\gamma})$  ( $= S^p$ ) while  $LN\bar{\gamma}$  birth take place at home, i.e. the entire population has been renewed, however part of any descendant generation are born abroad by migrants.

**Proposition 3.** *Migration flows over any time interval of length 1 are*

$$a) \text{ gross emigration: } M^e = LN\bar{\gamma}(1 - \underline{\gamma}) = \frac{LN\alpha y^*(y^* - y)}{y(y^* - \alpha y)}.$$

$$b) \text{ gross return migration of home born agents: } M^{r,h} = LN\bar{\gamma}(\bar{\gamma} - \underline{\gamma}) \\ = \frac{LN\alpha y^*(y^* - y)^2}{y(y^* - \alpha y)^2}.$$

$$\begin{aligned}
c) \text{ gross return migration of foreign born agents: } M^{r,f} &= LN(1 - \bar{\gamma}) \bar{\gamma} \\
&= \frac{LN\alpha y^*(y^* - y)(1 - \alpha)}{(y^* - \alpha y)^2}.
\end{aligned}$$

$$d) \text{ net migration: } M = M^e - M^{r,h} - M^{r,f} = 0$$

*Proof (sketch):* Since all individuals that migrate do so at time  $t_i = 0$ , proposition 3 a) is simply the sum of all home born ( $LN\bar{\gamma}$ ) temporary and permanent migrants. Proposition 3 b) follows from the fact that all cohorts behave identically, such that in equilibrium for every home born temporary outmigrant there is a matching temporary return migrant born in one of the previous cohorts. The return flow of foreign born agents in proposition 3 c) is composed of the share of stayers and temporary migrants ( $\bar{\gamma}$ ) in the total foreign born population ( $LN(1 - \bar{\gamma})$ ); or put differently, except for those agents that decide to be permanent migrants all other foreign born individuals will return at some point in time during their life to the home land. Proposition 3 d) follows from a), b) and c).

That the net flow of migrants associated with a given income differential and equilibrium stocks turns out to be zero is driven by the assumption that reproduction takes place at the end of an agents life time. If instead, one assumed that reproduction takes place earlier in an agents life, then some of the temporary migrants would reproduce while staying abroad generating an additional – and unmatched – flow of return migrants that results in *negative* net migration.

Figure 3 plots the migration flows given in proposition 3, for various levels of foreign income, when  $y = 100$ ,  $\alpha = 0.3$ ,  $N = 1000$  and  $L = 100$ .

**Corollary 1.** *All migration stocks  $S^p$ ,  $S^t$  and  $S$  and the migration flows  $M^e$ ,  $M^{r,h}$  and  $M^{r,f}$  are positive and increasing in the income differential  $y^* - y$ .*

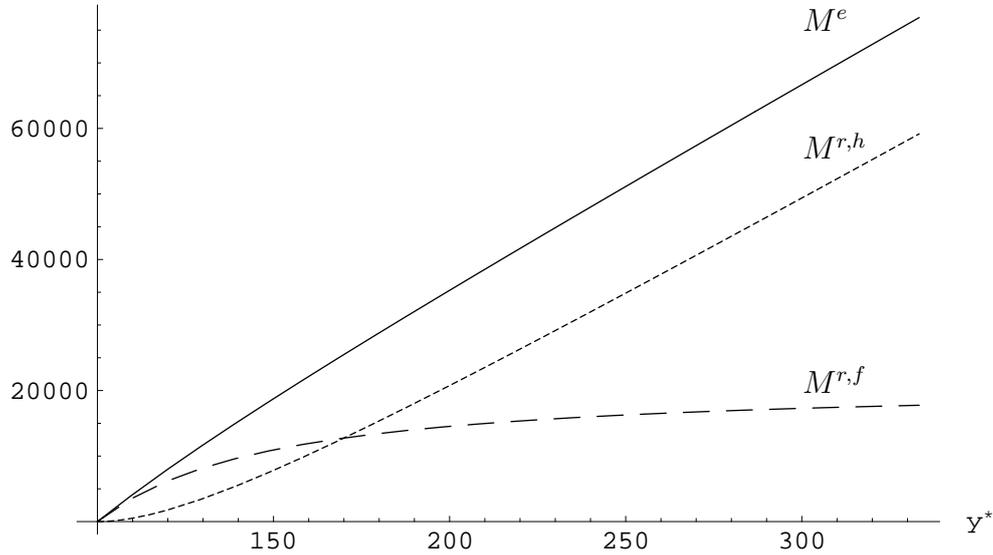


Figure 3: Migration flows in equilibrium

*The net migration flow  $M$  is zero and independent of the income differential  $y^* - y$ .*

### 3 Stock vs. flow models

The above results and particular Corollary 1 have important consequences for the empirical estimation of macro migration models. It follows from our model that an equilibrium relationship between the income differential and migration *stocks* instead of *net flows* emerges in equilibrium. A positive net migration flow can only occur during the transition to a steady state.<sup>11</sup> Yet, most macro migration models in the empirical literature to date state explicitly or implicitly that an equilibrium between migration *flows* and ex-

<sup>11</sup>Notice that this reasoning abstracts from differences in population growth rates among the migrant and the home population and other aspects such as the assimilation and naturalisation of migrants.

planatory variables such as the income differential exists, see e.g. the seminal contributions by Hatton (1995), Mayda (2010), and Pederson, Pytlikova, and Smith (2008). With the notable exception of (Hatton, 1995) this literature does not consider the time-series properties of the data.

In contrast, we examine the competing hypotheses that an equilibrium between the explanatory variables and migration stocks or migration flows exists empirically within a cointegration framework here. The concept of cointegration is closely related to the notion of equilibrium: a cointegration relationship between variables implies that economic forces drive the system towards the equilibrium defined by the long-run relationship posited (Engle and Granger, 1987). When considering long-run relationships, it becomes necessary to consider the underlying properties of the processes that generate time series variables. If variables follow different stochastic processes over time, spurious regression results can arise that suggest statistically significant long-run relationships between variables, when in fact this is merely evidence of contemporaneous correlations rather than meaningful causal relationships (Granger and Newbold, 1974).

Following the Engle and Granger (1987) procedure, we first test for both the stock and the flow model whether the dependent and the explanatory variables are integrated of the same order, and, if this is the case, whether the null hypothesis of no cointegration is rejected by our data at hand. We apply panel unit-root and panel cointegration tests, which increases the statistical power of the tests significantly in comparison to tests based on individual time series. Given a relatively short time dimension in our data set, we use in the final step an error correction specification for the estimation of the cointegrating vectors and the short-run dynamics. Based on this error correction specification we apply the test by Westerlund (2007) for drawing

inference on panel cointegration.

### 3.1 Two alternative specifications

The stock model of macro migration implied by Proposition 3, gives for the aggregate migration stock

$$mst = \alpha \ln \left( \frac{y^*}{(1-\alpha)y} - \frac{\alpha}{1-\alpha} \right),$$

where  $mst$  is defined as the share of the migration stock in the total population of the sending country, i.e.  $mst \equiv S/(LN)$ . For empirical purposes, one can approximate the expression for  $mst$  by

$$\ln mst = \beta_0 + \beta_1 \ln \left( \frac{y^*}{y} \right) + \beta_2 \ln(y),$$

which allows interpreting the estimated parameters as elasticities. We follow furthermore Todaro (1969) and Harris and Todaro (1970) in assuming that income levels are conditioned by employment opportunities in the respective locations. More specifically, if jobs are allocated in each period randomly among the workforce, we can write expected income as the wage times the employment rate, i.e. as  $w \times e$ . If individuals are risk averse and uncertainty focusses on employment opportunities, it can be expected that the coefficients for the employment variables are larger than those for the wage variables (Hatton, 1995). Moreover, since employment opportunities of migrants in host countries are below those of natives, the coefficient for the employment rate in the host country is larger than that in the source country. Finally, if capital markets are not perfect, liquidity constraints affect migration decisions. Consequently, for a given income difference between the host and the home country, the income level in the source country has a

positive impact on migration (Faini and Venturini, 1995).

Based on these considerations we derive the following parsimonious specification for the long-run migration function:

$$\ln mst_{it} = a_0 + a_1 \ln \frac{w_{ft}}{w_{it}} + a_2 \ln w_{it} + a_3 \ln e_{ft} + a_4 \ln e_{it} + \eta' \mathbf{x}_{fit} + \nu_{it}, \quad (17)$$

where  $mst_{it}$  denotes the migration stock as a percentage of the population in sending country  $i$ ,  $w_{ft}$  the wage rate in the host country,  $w_{it}$  the wage rate in the home country  $i$ ,  $e_{ft}$  the employment rate in the host country,  $e_{it}$  the employment rate in the home country  $i$ ,  $\mathbf{x}_{fit}$  a vector of variables which cover migration restrictions and other push and pull factors relevant for migration between sending country  $i$  and destination  $f$ , and  $\nu_{it}$  is the error term. The error term is specified as a one-way error component model (Hsiao, 1986), i.e. as  $\nu_{it} = \mu_i + \varepsilon_{it}$ , where  $\mu_i$  is a country-specific fixed effect and  $\varepsilon_{it}$  is white noise. Note that the country specific fixed effects cover all time invariant factors relevant for migration such as geographical distance, language proximity and long-term historical ties such as colonial links. Note that Brücker and Siliverstovs (2006) have proved that simple fixed effects estimators outperform heterogeneous and other more sophisticated estimators in the context of international migration. Finally,  $i = 1, \dots, N$  and  $t = 1, \dots, T$  denote the source country and time indices and  $f$  the destination country (Germany).

Compare this to the standard flow model of international migration. The flow equation in the empirical literature has the following form, see e.g. Hatton (1995), Mayda (2010) and Pederson, Pytlikova, and Smith (2008):

$$\begin{aligned} \ln m_{it} &= b_0 + b_1 \ln \frac{w_{ft}}{w_{it}} + b_2 \ln w_{it} + b_3 \ln e_{ft} + b_4 \ln e_{it} \\ &+ b_5 \ln mst_{it} + \lambda' \mathbf{x}_{fit} + \nu_{it}, \end{aligned} \quad (18)$$

where  $m_{it}$  denotes the net (gross) migration rate as a percentage of the population in sending country  $i$ . The existing stock of migrants is included on the right hand side of the model as a proxy for 'social network' effects which are expected to increase the propensity to migrate by alleviating the adaptation costs in the host country, see Hugo (1981), Massey and Espana (1987), Massey (1990a), Massey (1990b) and Bauer (1995).<sup>12</sup>

The estimation of the migration functions in equations (17) and (18) can be affected by spurious correlation effects if the regressions involve variables that follow an I(1) or other non-stationary process (see the seminal paper by Granger and Newbold (1974)). The notable exception is the situation when I(1) dependent and explanatory variables form a cointegration set, see Engle and Granger (1987). While there is a general agreement that macroeconomic variables such as income levels and employment rates are rather well represented as I(1) processes, there is still limited evidence on the time series properties of the migration flows and corresponding migrant stock variables.<sup>13</sup>

Particularly puzzling is the fact that the migration flow and the migration stock variable are included in equation (18). Since migration flows can be conceived as (almost) the first difference of migration stocks, they can hardly be I(1) variables if migration stocks are supposed to be I(1) variables as well. Thus, it is reasonable to expect that the migration flow variable is better approximated by an I(0) process if migration stocks are I(1). In this case it

---

<sup>12</sup>Of course, there exist more possible macro models of migration. Beyond the logarithmic functional form semi-logarithmic functions are also common in the literature (Hatton, 1995). Furthermore, some models use also measures for the inequality of earnings and further explanatory variables.

<sup>13</sup>There are few exceptions in the literature: Hatton (1995) provides empirical evidence that all variables in equation (18) are I(1) for UK-US migration from 1870 to 1913, but it is unclear whether this is also supported by other data sets. Brücker, Fachin, and Venturini (2011) and Fachin (2007) present mixed evidence for internal migration in Italy, while the findings of Brücker (2002) suggest that migration stocks in Germany contain a unit-root.

is suitable to use the stock model in equation (17) for estimating the long-run migration function.

### 3.2 Data

A cointegration analysis of the economic forces which drive international migration requires large time series of migration data. We use data from Germany, which is after the US and Russia the most important destination of migrants in the world in absolute terms. Migration stock and flow data by country of origin exist in Germany since 1967. Our data set covers the period 1967 to 2009, which gives 43 time series observations. We include 51 sending countries for which continuous time series were available. We distinguish a European and a non-European sample here. The European sample consists of 24 countries, i.e. the remaining 14 countries of the fifteen Member States of the European Union before 2004, Norway, Switzerland, Turkey, Bulgaria, the (former) Czechoslovakia, Hungary, Poland, Romania, the (former) USSR and the (former) Yugoslavia. For getting long-time series, we have aggregated the figures from the successor states of the former Czechoslovakia, USSR and Yugoslavia. The non-European sample contains Australia, the USA and Japan, 9 countries from the Middle East and Northern Africa, 9 countries from Asia (incl. China and India) and 5 countries from Sub-Saharan Africa.

The data on migration stocks and flows come from the German Federal Statistical Office (Statistisches Bundesamt, 2010a,b). For the stock of migrants, we use foreign residents as reported by the Central Register of Foreigners ('Ausländerzentralregister') as a variable.<sup>14</sup> The stock of foreign residents is reported on December 31 of each year (in some early years on

---

<sup>14</sup>Note that all residents of Germany are obliged to register their place of residence. The figures from the central register of foreigners are based on the reports of the municipalities.

September 30).<sup>15</sup> The number of foreign residents is slightly overstated by the Central Register of Foreigners, since return migration is not completely registered by the municipalities. Consequently, the figures for the stock of foreign residents has been revised following population censuses in 1972, 1987 and 2004. We include dummy variables which control for these breaks in the regressions if they appear significant.

The data on migration flows stem again from the central register of foreigners. We consider three flow variables: net migration flows,  $m_{it}$ , gross inflows,  $im_{it}$ , and gross return flows,  $out_{it}$ . The migration stock and flow variables are normalised by the population of the home countries, i.e. they are calculated as a percentage of the corresponding home population.

Population figures are depicted from the World Development Indicators 2010 (World Bank, 2010). Missing data has been complemented by OECD (2010) and sending country sources. As a proxy for wages and other incomes, the historical series of GDP per worker in purchasing power parities from the Penn World Tables 6.3 (Heston, Summers, and Aten, 2009) have been used and extrapolated by the growth rates from similar series in the World Development Indicators 2010. The employment rate is defined as 100 minus the unemployment rate in per cent. Unemployment rates have been

---

<sup>15</sup>It is sometimes argued that natural population growth and naturalisations distort the migration stock variable. By definition, the increase of the stock of foreign residents equals net immigration plus natural population growth minus the number of naturalisations plus reporting errors for a given period of time. Since our migration variables are calculated as rates, natural population growth cancels out if the rate of natural population growth of migrants equals the rate of natural population growth in the home countries. Thus, if natural population growth of the migrant population in Germany and the source country is similar and the rate of naturalisations is low, the annual increase of the stock of foreign residents as a share in the home population equals almost annual net immigration. Indeed, the annual increase of migration stocks as a share of home population almost equals net immigration rates in our sample. The German migration legislation has eased the naturalisation of immigrants in 1999. As a robustness check, we included a dummy variable which has a value of 1 from 1999 onwards and zero otherwise, which however did not turn out to be significant.

taken from the OECD Stat database (OECD, 2010) and complemented by the World Development Indicators 2010. The ILO definition has been used for all unemployment rates. Particularly in the non-European sample, only few countries report complete (un-)employment series. Considering the sending countries' employment rate reduces thus our sample from 2193 to 1379 observations. In our standard specification, we do therefore not consider the sending countries' employment rate.

Institutions and other push and pull factors affect migration decisions in one way or another. We consider therefore in addition to our basic economic variables the following dummy variables: A free movement dummy ( $FREE_{it}$ ), which has a value of 1 if the rules of free movement within the European Union and European Economic Area apply for country  $i$  and of zero if otherwise, a dummy which covers the emigration restrictions in the former Eastern bloc ( $IRON_{it}$ ), a dummy for dictatorship in country  $i$  ( $DICT_{it}$ ) which has been derived from the Freedom House political rights index, a dummy for civil and other wars in sending country  $i$ , and finally a dummy variable which covers the change of asylum legislation in Germany which became effective in the year 1992 ( $ASYL_{ft}$ ).

The list of variables and the descriptive statistics are presented in Annex Tables A1 and A2.

### 3.3 Testing for unit roots

In the first step of the empirical analysis, the variables are tested for unit-roots for drawing inference on the order of integration. To this end, we apply the panel-unit root test suggested by Breitung (2000). The null hypothesis is that all panels contain a unit-root and the alternative that all series are stationary. Although the Breitung (2000) test assumes that all panels have

a common autoregressive parameter, Breitung and Das (2005) provide evidence that the test also has considerable power in the heterogeneous case where each panel is allowed to have its own autoregressive parameter. Monte Carlo simulations suggest that the test is substantially more powerful than other panel unit-root tests for data sets of modest size. Importantly, there is an option for the Breitung (2000) test which is robust for cross-sectional correlation of the error terms. This statistic has an asymptotically normal distribution when first the time dimension tends to infinity followed by the group dimension tending to infinity.

The cross-sectional correlation robust version of the Breitung (2000) test requires that  $T > N$ , which is not the case in our total sample with 51 countries and 43 time periods. We apply therefore first the non-robust version of the panel unit-root test to the total sample. Following Levin, Lin, and Chu (2002) we subtracted the mean of the series across panels from the series in order to mitigate the impact of cross-sectional dependence. In the European and the non-European samples, where the group dimension is smaller than the time dimension of the panel, we used the version of the panel unit-root test which is robust to cross-sectional correlation of the error terms.<sup>16</sup>

The Breitung (2000) test requests that the panel is strictly balanced. That holds for all variables except for the employment rate in the sending countries. For this variable we applied the Im, Pesaran, and Shin (2003) (IPS-test) for panel unit-roots. This tests allows each panel to have its own autoregressive parameter and does not require that the data are strongly balanced. The null hypothesis is that all panels have a unit-root. The alternative hypothesis is that the fraction of panels that are stationary is nonzero. We report the  $\bar{z}_t$ -statistics, which assumes that the group dimension tends to

---

<sup>16</sup>Our findings turn out to be robust also in other compositions of the sample not presented here.

infinity. Finally, for the employment rate in Germany, a standard Augmented Dickey-Fuller (ADF) test is used.

---

Table 1 about here

---

Table 1 present the results of the unit-root tests performed on the host- and home-country-specific economic variables. As expected, the null hypothesis that the time-series of the macroeconomic variables, i.e. the ratio of the GDP per worker, the GDP per worker in the sending country<sup>17</sup> and the employment rates contain a unit-root, is not rejected by the panel unit root tests. In contrast, for the first difference of these variables, the null hypothesis of a unit-root is rejected at the 1%-level in all panel unit-root tests. While the ADF-test suggests that the German employment variable contains a unit-root if we do not consider a trend variable, the test results indicate that this variable might be trend-stationary. However, since a deterministic time trend in the evolution (un-)employment rates does not make sense for a priori reasons, we think that the latter result is not relevant for our further investigation.

For the migration stock variable, the panel unit-root tests do not reject the null hypothesis of a unit-root, while this hypothesis is rejected for the net migration rate. Note that we cannot test the net migration rate in logarithmic form, since there are many negative observations. The results for the gross immigration and the gross return migration rates are mixed: While the null of a unit-root is rejected for the gross immigration rate in the total sample

---

<sup>17</sup>Note that the unit-root tests obtain the same results for the ratio of the income variables and for the income variables in the sending countries since the means of the series across panels are subtracted in order to mitigate the impact of cross-sectional dependence. We report the results only for completeness.

and the non-European sample, it is not rejected in the European sample. The null of a unit-root is not rejected for the return migration rate in the total and the European sample, but rejected in the non-European sample.

Thus, we can conclude from the panel unit-root tests that the assumption of the standard migration model, i.e. that net migration *flows* on the one hand, and macroeconomic variables such as GDP per worker levels or employment rates on the other hand, are integrated of the same order, is not supported by the data set employed here. As a consequence, the net flow variable and the explanatory variables cannot form a cointegrated set. However, we find strong evidence that the migration stock and the explanatory variables are  $I(0)$  variables. The evidence for the gross immigration and return variables is mixed, such that we cannot draw a clear-cut conclusion here. Note that our theoretical model predicts that the gross immigration and return migration rates are a fraction of the migration stock in equilibrium, such that they might be cointegrated with the explanatory variables similar to the migration stock variable. However, empirically we find only mixed evidence that the data-generating processes of the gross flow and the explanatory variables have the same stochastic properties, such that we do not consider the gross flow variables in the estimation of the cointegration vectors.

### **3.4 Testing for panel cointegration**

In the next step we test whether the migration stock and the explanatory variables indeed form a cointegrated set. To this end we apply the procedure suggested by Westerlund (2007), which tests the null hypothesis of no cointegration by determining whether the individual panel members are error correcting. The tests are very flexible and allow for an almost

completely heterogeneous specification of both the long- and short-run parts of the error-correction model, where the latter can be determined from the data. The series are allowed to be of unequal length.

Table 2 displays the group mean statistics,  $G_{\tau}$ -, and the panel  $P_{\tau}$ -statistics. The group-mean statistics start from a weighted average of the individually estimated error correction parameters and their  $t$ -statistics, respectively. It tests the null hypothesis that the error correction parameter is zero for all  $i$  versus the alternative that the error correction parameter is smaller than zero for at least one group  $i$ . The panel  $P_{\tau}$  test statistics pool the information over the cross-sectional units to test the null hypothesis that error correction parameter is zero for all  $i$  versus the alternative hypothesis that it is smaller than zero for all  $i$ . Thus, a failure to reject the null hypothesis should be taken as a failure to reject the null of no cointegration for the panel as a whole in the latter case.

We test two models for cointegration: The first model explains the migration stock by the ratio of the GDP per worker, the GDP per worker in the sending country and the employment rate in Germany. The second model includes also the employment rate of the sending countries. Note that the second model is an unbalanced panel. The lag structure of the error correction model is chosen by the Akaike information criteria. We consider between 1 and 3 lags of the first differences of the variables in the balanced panel model, in the unbalanced panel we were only able to consider one lag.

---

Table 2 about here

---

In the balanced panel, the null hypothesis of no cointegration is rejected both by the group mean and the panel statistics in the total sample and in

the European sample. While the group mean statistics fails to reject the null of no cointegration in the non-European sample, the panel cointegration test statistics rejects the null of no cointegration for the panel as a whole. In the unbalanced panel, the picture is more mixed: While the panel cointegration test statistics rejects the null hypothesis of no cointegration for the total and the European sample, the group mean statistics rejects the null of no cointegration only in the European sample.

Thus, the results of the cointegration tests suggest that we cannot reject the hypothesis that the variables of the stock model form a cointegrated set in the balanced panel, although we have to take this finding with a grain of salt for the non-European countries where migration episodes fluctuate more over time than in the European context. However, if we include employment rates of the sending countries, our results turn out to be much weaker. This might be also traced back to the fact that the panel is not balanced any more and that we lose particularly in the non-European sample many time-series observations.

## 4 Estimation results

There are different procedures for estimating both the long-run cointegration relationship and the short-run dynamics. If the variables form a cointegrated set, the cointegrating vector can be consistently estimated in a static regression which completely omits the dynamics of the model (Engle and Granger, 1987). Although the famous super-consistency result (Stock, 1987) indicates that convergence is rather fast, the asymptotic distribution of the least squares estimator and the associated  $t$ -statistics is non-normal in finite samples. Moreover, an unaddressed 'endogeneity bias' invalidates standard

hypothesis testing in samples of finite size.<sup>18</sup> Monte-Carlo evidence suggests that the estimation bias of the cointegrating parameter is smaller in dynamic than in static models (Banerjee, Dolado, Henry, and Smith, 1986). The empirical equation is therefore specified here in form of an error correction model (ECM), which allows estimation of both the long-term cointegrating vector and the short-run dynamics. Note that the ECM has a flexible functional form and imposes few restrictions on the adjustment process.

Specifically, the estimation model has the form

$$\begin{aligned} \Delta \ln mst_{it} = & \beta_1 \ln mst_{i,t-1} + \beta_2 \ln \frac{w_{f,t-1}}{w_{i,t-1}} + \beta_3 \ln w_{i,t-1} + \beta_4 \ln e_{f,t-1} + \\ & \beta_5 \ln e_{i,t-1} + \beta_6 \Delta \ln \frac{w_{ft}}{w_{it}} + \beta_7 \Delta \ln w_{it} + \beta_8 \Delta \ln e_{ft} + \\ & \beta_9 \Delta \ln e_{it} + \beta_{10} \Delta \ln mst_{i,t-1} + \eta' \mathbf{z}_{it} + \mu_i^* + \varepsilon_{it}, \end{aligned} \quad (19)$$

where  $\mu_i^* = \mu_i / -\beta_1$  is the long-run value for the country-specific effect,  $\Delta$  the first-difference operator,  $\mathbf{z}_{it}$  a vector of institutional variables and  $\eta$  the corresponding vector of coefficients. As outlined above, we use dummy variables which should capture institutions and push and pull factors such as dictatorship and wars affecting migration incentives and opportunities. Note that the adjustment parameter of the ECM is given by  $-\beta_1$ , and that the long-term coefficients of the cointegrating relationship are given by  $-\beta_k / \beta_1$ , where  $k = 2, 3, \dots, 5$ . Further lags of the first differences of the dependent variable and lags of first differences of the explanatory variables have not been considered in this specification of the ECM since they appear not significant.

The estimation results of the short-run semi-elasticities of the dynamic model are presented in Tables 3 and 4. The model is estimated with a standard fixed effects (within) model (FE) using the log of the population

<sup>18</sup>See Patterson (2000) for a detailed discussion.

size as a weight. The results of the  $F$ -test show that the country-specific effects are significant at the 1%-level in all regressions.

---

Table 3 about here

---

Table 3 presents the results for the total sample. We have estimated three models: The first model considers the ratio of the GDP per worker, the GDP per worker in the sending country and the German employment rate as explanatory variables. The second model considers also the dummy variables for institutions and other push and pull factors. Finally, the third model includes the employment rate of the sending countries in addition.

In all three regressions the coefficient of the lagged migration stock variable is significantly negative, which assures the dynamic stability of the model. However, the coefficient on the lagged first difference of the migration stock has a value of about 0.5 and is highly significant, which might be interpreted as a hint that 'chain' or 'herd' effects (Epstein and Hillman, 1998) play an important role in the short-term – although they do not affect the dynamic stability of the system in the long-run.

All coefficients on the economic variables have the expected signs and appear highly significant: Both the log difference of the GDP per worker and the GDP per worker in the sending countries exert a positive and highly significant impact on migration stocks. While the first finding suggests that income differences have a substantial impact on migration, the latter one suggests that tends to increase with the income level in the sending countries at a given income difference. This supports previous findings in the literature and suggests that liquidity constraints may play an important role

for migration decisions (Faini and Venturini, 1995). The German employment rate is highly significant indicating that employment opportunities in the destination countries have a substantial impact on migration. Finally, the employment rate in the sending countries has the expected negative coefficient and appears highly significant as well. However, this parameter has a much smaller value than the employment rate in the receiving country. Note that many empirical studies find that home employment opportunities have been insignificant or have actually increased migration (see Greenwood (1975) for a review). Our results here do not confirm these findings, but they do show that home employment has a much weaker impact than employment in the receiving countries. One possible explanation for this phenomenon is that favorable employment opportunities in home countries might have ambiguous effects on migration, since higher employment rates reduce incentives to seek employment abroad on the one hand, while on the other they help to lift liquidity constraints that potential migrants face, and in doing so, may encourage migration.

The institutional variables have the expected signs in all regressions, but we observe substantial differences in the size of the parameters and their significance. The coefficient of the free movement dummy is only significant in the balanced panel, while it appears insignificant in the regression which considers the unemployment rate of the sending countries. Not surprisingly, the negative coefficient on the iron curtain dummy indicates that the emigration restrictions in the countries of the former Eastern Bloc have effectively hindered migration from there to Germany. The dictatorship and the war dummies have the expected signs, but are not significant in both specifications of the model. However, the restrictions on asylum which became effective in 1992 have significantly hindered migration to Germany.

---

Table 4 about here

---

Our findings from the European- and the non-European samples confirm this picture: In all regressions the negative coefficient on the lagged migration stock indicates that the model is stable from a dynamic perspective. However, the large and highly significant coefficient on the first difference of the migration stock variables suggests that the chain- or network effects play an important role. Interestingly enough, the coefficient on the income ratio is not only highly significant in both samples, it also has a similar size. However, the size of the coefficient on the income variable in the sending countries is larger in the non-European sample suggesting that liquidity constraints are more important there than in the European sample. The German employment rate exerts a large and highly significant effect on migration stocks in all four regressions. The employment rate of the sending countries affects migration stocks negatively, but the size of the coefficient is much smaller compared to that of the employment rate in Germany. The coefficients of the institutional variables have the expected signs, but appear not always significant: As an example, the free movement dummy is not significant at all, while the iron curtain dummy appears only significant in the balanced dummy which does not consider the sending countries' employment rate. In contrast, the dictatorship, war and asylum dummies appear significant in the first three regressions. Note that the last regression is based only on a limited number of observations, since many non-European countries do not report their employment rates.

All together, our regression results tend to confirm our theoretical expectations and particularly the economic variables appear significant in all

regressions. Moreover, the size of the coefficients is also relatively homogeneous in most regressions which supports the application of the standard fixed effects estimator with uniform slope coefficients as suggested by Brücker and Siliverstovs (2006).

## 5 Conclusion

In this paper, we examined the macro determinants of migration both from a theoretical and an empirical perspective. The model presented in the theoretical part establishes a long-run equilibrium, in which individuals can stay their entire life in the home country, migrate temporarily abroad – with individual durations of migration spells – or stay permanently in a foreign country depending on their preferences. This model generated insights in the mechanics of migration stocks and flows. The number of migrants, the duration of migration spells and therewith the stock of migrants all increases with the income difference between the host and the home country, while net migration ceases to zero. The gross emigration and return migration rates are related to the stocks of permanent and temporary migrants. Since the stock of permanent and temporary migrants are a positive function of the income differential. Consequently, existing empirical migration models, estimating net migration flows, instead of stocks, may be misspecified.

In the empirical part of the paper the determinants of international migration have been analysed within a cointegration framework. The methodological aspects of the analysis can be summarised as follows: first, the results of the panel unit-root and panel cointegration test suggest that the standard net flow migration model is misspecified – at least for the data set used here. The traditional migration model in the empirical model explains migration

flows by a number of explanatory variables such as GDP per capita, (un-)employment rates, (lagged) migration stocks and institutional variables. It is widely acknowledged in the literature that macroeconomic variables such as GDP and employment are non-stationary variables, or, more specifically,  $I(1)$  variables. The existence of a long-run equilibrium between migration flows and the traditional set of macroeconomic variables requires therefore that migration flows are  $I(1)$  as well. The tests carried out in the empirical part of this paper, however, indicate that migration rates are stationary, while migration stocks are  $I(1)$  variables. Moreover, our test results reject the null hypothesis of no cointegration for most specifications of our error correction model. This can be interpreted as empirical support for the theoretical hypothesis that migration stocks and explanatory variables such as the income differential and employment variables form an equilibrium relationship.

Our findings have some important policy consequences. The flow model suggests that migration does not stop before expected income levels between host and source countries have converged to a certain threshold level, which is determined by the costs of migration. In case of persistent differences in expected income levels, either the total population will eventually migrate or migration will not occur in the first place. In contrast, the stock model predicts that migration ceases when the benefits of migration equal the costs to the marginal migrant, such that a long-run equilibrium between migration stocks and expected income emerges. Consequently, migration may cease despite the existence of large income differences.

## References

- BANERJEE, A., J. J. DOLADO, D. F. HENRY, AND G. W. SMITH (1986): “Exploring Equilibrium Relationships in Econometrics through Static Models : Some Monte Carlo Evidence,” *Oxford Bulletin of Economics and Statistics*, 48(3), 253 – 277.
- BAUER, T. (1995): “The migration decision with uncertain costs,” *Mnchner Wirtschaftswissenschaftliche Beitrge* 95-25, University of Munich.
- BEINE, M., F. DOCQUIER, AND ÇAGLAR ÖZDEN (2011): “Diasporas,” *Journal of Development Economics*.
- BREITUNG, J. (2000): “The local power of some unit root tests for panel data,” in *Advances in Econometrics, Volume 15: Nonstationary Panels, Panel Cointegration, and Dynamic Panels*, ed. by B. H. Baltagi, pp. 161 – 178. JAI Press, Amsterdam.
- BREITUNG, J., AND S. DAS (2005): “Panel unit root tests under cross-sectional dependence,” *Statistica Neerlandica*, 59, 414 – 433.
- BRÜCKER, H. (2002): “Can International Migration Solve the Problems of European Labour Markets,” *Economic Survey of Europe*, 2002/2, 109 – 142.
- BRÜCKER, H., S. FACHIN, AND A. VENTURINI (2011): “Do Foreigners Replace Native Immigrants? Evidence From a Time-Series Analysis in Italy,” *Economic Modelling*, (forthcoming), 109 – 142.
- BRÜCKER, H., AND B. SILIVERSTOVVS (2006): “On the estimation and forecasting of international migration: how relevant is heterogeneity across countries?,” *Empirical Economics*, 31(3), 735–754.

- DJAJIC, S., AND R. MILBOURNE (1988): "A General Equilibrium Model of Guestworker Migration," *Journal of International Economics*, 25, 335–351.
- DUSTMANN, C. (1995): "Return migration - the European experience," *Economic Policy*, 22, 214–250.
- (2003): "Return migration, wage differentials, and the optimal migration duration," *European Economic Review*, 47, 353–369.
- DUSTMANN, C., AND O. KIRCHKAMP (2002): "The Optimal Migration Duration and Activity Choice After Re-migration," *Journal of Development Economics*, 67(2), 351–372.
- ENGLE, R. F., AND C. W. J. GRANGER (1987): "Cointegration and Error Correction: Representation, Estimation and Testing," *Econometrica*, 55, 251–276.
- EPSTEIN, G. S., AND A. L. HILLMAN (1998): "Herd Effects and Migration," CEPR Discussion Paper: 1811.
- FACHIN, S. (2007): "Long-Run Trends in Internal Migrations in Italy: a Study in Panel Cointegration with Dependent Units," *Journal of Applied Econometrics*, pp. 401–428.
- FAINI, R., AND A. VENTURINI (1995): "Migration and Growth: The Experience of Southern Europe," CEPR Discussion Paper No. 964.
- FREEDOMHOUSE (2008): *Freedom in the World 2008: The Annual Survey of Political Rights and Civil Liberties*. Freedom House (available on-line at <http://www.freedomhouse.org/uploads>).
- GRANGER, C. W. J., AND P. NEWBOLD (1974): "Spurious Regression in Econometrics," *Journal of Econometrics*, 2(2), 111–120.

- GREENWOOD, M. J. (1975): "Research on Internal Migration in the United States: A Survey," *Journal of Economic Literature*, 13(2), 397–433.
- GROGGER, J., AND G. H. HANSON (2011): "Income maximization and the selection and sorting of international migrants," *Journal of Development Economics*, forthcoming (available online).
- HARRIS, J. R., AND M. TODARO (1970): "Migration unemployment and development: a two sector analysis," *American Economic Review*, 60, 126 – 142.
- HATTON, T. J. (1995): "A Model of U.K. Migration, 1870-1913," *Review of Economics and Statistics*, 77, 407–415.
- HESTON, A., R. SUMMERS, AND B. ATEN (2009): *Penn World Table Version 6.3*. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.
- HILL, J. K. (1987): "Immigrant decisions concerning the duration of stay and migration frequency," *Journal of Development Economics*, 25, 221–234.
- HSIAO, C. (1986): *Analysis of Panel Data*. Cambridge University Press, Cambridge MA.
- HUGO, G. (1981): "Village-Community Ties, Village Norms, and Ethnic and Social Networks: A Review of the Evidence from the Third World," in *Migration Decision Making: Multidisciplinary Approaches to Microlevel Studies in Developed and Developing Countries*, ed. by G. de Jong, and R. Gardener, pp. 186–225. Pergamon Press, New York.

- IM, K. S., M. PESARAN, AND Y. SHIN (2003): "Testing for Unit Roots in Heterogeneous Panels," *Journal of Econometrics*, 115, 53–74.
- KARRAS, G., AND C. U. CHISWICK (1999): "Macroeconomic Determinants of Migration: The Case of Germany 1964-1988," *International Migration*, 37(4), 657–677.
- LEVIN, A., C.-F. LIN, AND C.-S. J. CHU (2002): "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties," *Journal of Econometrics*, 108(1), 1–24.
- MASSEY, D. (1990a): "The Social and Economic Origins of Immigration," *Annals of the American Academy of Political and Social Sciences*, No. 510.
- MASSEY, D. S. (1990b): "Social Structure, Household Strategies and the Cumulative Causation of Migration," *Population Index*, 56, 1–26.
- MASSEY, D. S., AND F. ESPANA (1987): "The Social Process of International Migration," *Science*, 237, 733–738.
- MAYDA, A. M. (2010): "International migration: a panel data analysis of the determinants of bilateral migration flows," *Journal of Population Economics*, 23, 1249–1274.
- MESNARD, A. (2004): "Temporary migration and capital market imperfections," *Oxford Economic Papers*, 56, 242 – 262.
- OECD (2010): *OECD Statistical Database*. OECD online, Paris.
- PATTERSON, K. (2000): *An Introduction to Applied Econometrics: A Time Series Approach*. PALGRAVE, New York.

- PEDERSON, P. J., M. PYTLIKOVA, AND N. SMITH (2008): "Selection or Network Effects? Migration Flows Into 27 OECD Countries, 1990-2000," *European Economic Review*, 52(7), 1160–1186.
- RAVENSTEIN, E. G. (1889): "The Laws of Migration," *Journal of the Statistical Society*, 59, 214–301.
- STARK, O. (1995): "Return and Dynamics: The Path of Labor Migration when Workers Differ in their Skills and Information is Asymmetric," *Scandinavian Journal of Economics*, 97(1), 55–71.
- STATISTISCHESBUNDESAMT (2010a): "Ausländische Bevölkerung nach Statsangehörigkeit, special provision," Statistisches Bundesamt: Wiesbaden.
- (2010b): "Wanderungen ber die Außengrenzen Deutschlands nach Statsangehörigkeit, special provision," Statistisches Bundesamt: Wiesbaden.
- STOCK, J. (1987): "Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors," *Econometrica*, 55, 1035–1056.
- TODARO, M. (1969): "A model of labour migration and urban unemployment in less developed countries," *American Economic Review*, 59(1), 139–148.
- WESTERLUND, J. (2007): "Testing for error correction in panel data," *Oxford Bulletin of Economics and Statistics*, 69, 709–748.
- WORLD BANK (2010): *World Development Indicators 2010*. World Bank, Washington, DC.

## A Appendix

### A.1 First-order conditions of the Lagrangian of the migrant's maximisation problem

Define by  $L$  the Lagrangian for the maximisation problem in (3) under the budget constraint (4). The first-order conditions are:

$$\frac{\partial L}{\partial c_i^*(t)} = \alpha \gamma_i^{1-\alpha} c_i^{*-(1-\alpha)} - \lambda = 0, \quad (\text{A.1})$$

$$\frac{\partial L}{\partial c_i(t)} = \alpha c_i^{-(1-\alpha)} - \lambda = 0, \quad (\text{A.2})$$

$$\frac{\partial L}{\partial \tau_i} = \gamma_i^{1-\alpha} c_i^*(t)^\alpha + c_i(t)^\alpha - \lambda(c_i^*(t) - c_i(t) + y - y^*) = 0, \quad (\text{A.3})$$

$$\frac{\partial L}{\partial \lambda} = \tau_i y^* + (1 - \tau_i)y - \tau_i c_i^*(t) - (1 - \tau_i)c_i(t) = 0, \quad (\text{A.4})$$

where  $\lambda$  is the shadow value of wealth.

### A.2 Derivatives of the optimal $\tau_i$

Differentiating (11) with respect to  $y^*$ ,  $y$  and  $\gamma_i$  gives:

$$\frac{\partial \tau_i}{\partial y^*} = \frac{y - y\alpha}{(y^* - y)^2} > 0, \quad (\text{A.5})$$

$$\frac{\partial \tau_i}{\partial y} = \frac{y^* (\alpha - 1)}{(y^* - y)^2} < 0, \quad (\text{A.6})$$

$$\frac{\partial \tau_i}{\partial \gamma_i} = \frac{\alpha}{(1 - \gamma)^2} > 0. \quad (\text{A.7})$$

Table 1: Panel unit-root tests

|  | total sample  |             | European sample |             | non-European sample |             |
|--|---------------|-------------|-----------------|-------------|---------------------|-------------|
|  | without trend | with trend  | without trend   | with trend  | without trend       | with trend  |
| <i>Breitung t-statistics</i>                             |               |             |                 |             |                     |             |
| $\ln mst_{it}$   | 8.852         | 9.637       | 1.906           | 1.470       | 1.748               | 2.826       |
| $\Delta \ln mst_{it}$                                    | -12.237 ***   | -14.812 *** | -4.579 ***      | -5.275 ***  | -6.202 ***          | -8.259 ***  |
| $m_{it}$   | -10.160 ***   | -9.351 ***  | -2.440 ***      | -6.888 ***  | -4.036 ***          | -13.421 *** |
| $\ln im_{it}$  | -1.704 **     | -1.083      | 0.569           | 1.715       | -3.860 ***          | -5.259 ***  |
| $\Delta \ln im_{it}$                                     | -21.505 ***   | -21.814 *** | -5.614 ***      | -6.133 ***  | -15.250 ***         | -20.364 *** |
| $\ln out_{it}$   | 0.689         | -0.573      | 0.899           | 1.033       | -3.084 ***          | -3.482 ***  |
| $\Delta \ln out_{it}$                                    | -21.116 ***   | -23.699 *** | -4.661 ***      | -5.311 ***  | -11.747 ***         | -12.494 *** |
| $\ln \frac{w_{ft}}{w_{it}}$                              | 4.864         | 11.356      | 0.711           | 3.385       | 1.035               | -0.282      |
| $\Delta \ln \frac{w_{ft}}{w_{it}}$                       | -15.867 ***   | -14.016 *** | -8.882 ***      | -7.416 ***  | -9.193 ***          | -10.003 *** |
| $\ln w_{it}$   | 4.864         | 11.356      | 0.711           | 3.385       | 1.035               | -0.282      |
| $\Delta \ln w_{it}$                                      | -15.867 ***   | -14.016 *** | -8.882 ***      | -7.416 ***  | -9.193 ***          | -10.003 *** |
| <i>Im-Peseran-Shin <math>\bar{z}_t</math>-statistics</i> |               |             |                 |             |                     |             |
| $\ln e_{it}$   | 0.645         | 0.260       | 1.699           | 0.864       | -0.573              | 0.289       |
| $\Delta \ln e_{it}$                                      | -13.907 ***   | -14.504 *** | -10.222 ***     | -10.429 *** | -9.274 ***          | -10.167 *** |
| <i>Augmented Dickey-Fuller test statistics</i>           |               |             |                 |             |                     |             |
| $\ln e_{ft}$   | -1.719        | -4.187 **   | -               | -           | -                   | -           |
| $\Delta \ln e_{ft}$                                      | -4.852 ***    | -4.919 ***  | -               | -           | -                   | -           |

Notes: \*\*\*, \*\*, \* denote the 1%-, 5%- and 10%-significance levels, respectively.– The Breitung test-statistics is based on the standard t-statistics. In the total sample, in each time period the mean of the observations are subtracted from the series in order to mitigate the cross-sectional correlation of the error terms. In the European and the non-European samples the t-statistics which is robust to cross-sectional correlation of the error terms is reported. This statistics has an asymptotically normal distribution if first the time dimension of the panel tends to infinity followed by the group dimensions tending to infinity.– The Im-Peseran-Shin test reports the  $\bar{z}_t$ -tilde bar statistics which allow that the panel is unbalanced. The  $\bar{z}_t$ -tilde bar statistics assumes that the group dimension tends to infinity. The number of observations are sufficient for the asymptotic normal distribution to hold. In order to mitigate cross-sectional correlations of the error terms, the mean of the observations are subtracted from the series in each time period.– The significance levels of the Augmented Dickey Fuller Test are based on the MacKinnon critical values.

Table 2: Westerlund's test for panel cointegration

|  | <b>total sample</b> |             |     | <b>European sample</b> |             |     | <b>non-European sample</b> |             |    |
|--|---------------------|-------------|-----|------------------------|-------------|-----|----------------------------|-------------|----|
|  | Value               | Z-statistic |     | Value                  | Z-statistic |     | Value                      | Z-statistic |    |
| Model with GDP per worker and German employment rate       |                     |             |     |                        |             |     |                            |             |    |
| $G_\tau$   | -2.838              | -5.966      | *** | -3.626                 | -7.846      | *** | -2.138                     | -0.803      |    |
| $P_\tau$   | -18.498             | -5.487      | *** | -18.282                | -8.179      | *** | -10.531                    | -1.682      | ** |
| Model with GDP per worker, German and home employment rate |                     |             |     |                        |             |     |                            |             |    |
| $G_\tau$   | -2.373              | -1.115      |     | -2.919                 | -3.47       | *** | -1.537                     | 2.759       |    |
| $P_\tau$   | -16.545             | -3.561      | *** | -14.694                | -4.481      | *** | -5.101                     | 2.358       |    |

Table 3: Estimation results: total sample

|                                    | model (1) |     | model (2) |        | model (3) |       |        |     |       |
|------------------------------------|-----------|-----|-----------|--------|-----------|-------|--------|-----|-------|
|                                    | Coeff.    | SE  | Coeff.    | SE     | Coeff.    | SE    |        |     |       |
| $\ln mst_{i,t-1}$                  | -0.012    | *** | 0.004     | -0.017 | ***       | 0.004 | -0.013 | **  | 0.005 |
| $\ln \frac{w_{ft}}{w_{it}}$        | 0.231     | *** | 0.055     | 0.367  | ***       | 0.057 | 0.414  | *** | 0.051 |
| $\ln w_{i,t-1}$                    | 0.172     | *** | 0.054     | 0.314  | ***       | 0.057 | 0.324  | *** | 0.049 |
| $\ln e_{f,t-1}$                    | 2.372     | *** | 0.233     | 2.393  | ***       | 0.230 | 2.293  | *** | 0.196 |
| $\ln e_{i,t-1}$                    |           |     |           |        |           |       | -0.250 | *** | 0.066 |
| $FREE_{i,t-1}$                     |           |     |           | 0.034  | **        | 0.017 | 0.014  |     | 0.012 |
| $IRON_{i,t-1}$                     |           |     |           | -0.111 | ***       | 0.019 | -0.060 | *** | 0.020 |
| $DICT_{i,t-1}$                     |           |     |           | 0.021  |           | 0.015 | 0.099  | *** | 0.015 |
| $WAR_{i,t-1}$                      |           |     |           | 0.134  | ***       | 0.034 | -0.036 |     | 0.035 |
| $ASYL_{i,t-1}$                     |           |     |           | -0.058 | ***       | 0.009 | -0.033 | *** | 0.007 |
| $\Delta \ln mst_{i,t-1}$           | 0.479     | *** | 0.017     | 0.481  | ***       | 0.017 | 0.502  | *** | 0.020 |
| $\Delta \ln \frac{w_{ft}}{w_{it}}$ | 0.089     |     | 0.216     | 0.441  | **        | 0.219 | 1.171  | *** | 0.187 |
| $\Delta \ln w_{it}$                | -0.086    |     | 0.217     | 0.318  |           | 0.221 | 0.643  | *** | 0.188 |
| $\Delta \ln e_{ft}$                | 0.804     | *** | 0.273     | 0.902  | ***       | 0.272 | 1.146  | *** | 0.226 |
| $\Delta \ln e_{it}$                |           |     |           |        |           |       | -0.220 |     | 0.134 |
| $\Delta FREE_{it}$                 |           |     |           | 0.023  |           | 0.047 | 0.012  |     | 0.031 |
| $\Delta IRON_{it}$                 |           |     |           | -0.276 | ***       | 0.060 | -0.323 | *** | 0.048 |
| $\Delta DICT_{it}$                 |           |     |           | 0.055  |           | 0.037 | 0.100  | *** | 0.028 |
| $\Delta WAR_{it}$                  |           |     |           | 0.100  | ***       | 0.030 | -0.009 |     | 0.013 |
| $\Delta ASYL_{ft}$                 |           |     |           | 0.011  |           | 0.016 | 0.040  | *** | 0.003 |
| obs.                               | 2091      |     |           | 2091   |           |       | 1319   |     |       |
| groups                             | 51        |     |           | 51     |           |       | 41     |     |       |
| balanced panel                     | yes       |     |           | yes    |           |       | no     |     |       |
| $R^2$                              | 0.38      |     |           | 0.42   |           |       | 0.50   |     |       |

Notes: \*\*\*, \*\*, \* denote the 1%-, 5%- and 10% significance levels, respectively.– The dependent variable is the first difference of the log of the migration stock,  $\Delta \ln mst_{it}$  normalized by the sending country population.– All regressions are estimated by weighted OLS, using the log of the population size as weights.–All Regressions include a dummy variable which controls for a statistical break in the time series in 1987.– The F(50, 2031) test statistic for country-specific fixed effects is 3.06\*\*\* in the first, the F(50,2021) test statistic is 3.77\*\*\* in the second, and the F(40, 1256) is 7.05\*\*\* in the third regression, which rejects the H0 of no country-specific fixed effects at the 1%-level.

Table 4: Regression results: European and non-European samples

|                                    | European sample |     |        |        | non-European sample |       |        |     |      |        |     |       |
|------------------------------------|-----------------|-----|--------|--------|---------------------|-------|--------|-----|------|--------|-----|-------|
|                                    | Coeff.          | SE  | Coeff. | SE     | Coeff.              | SE    | Coeff. | SE  |      |        |     |       |
| $\ln mst_{i,t-1}$                  | -0.013          | **  | 0.006  | -0.014 | *                   | 0.007 | -0.017 | *** | 0.01 | -0.037 | *** | 0.009 |
| $\ln \frac{w_{ft}}{w_{it}}$        | 0.436           | *** | 0.077  | 0.421  | ***                 | 0.069 | 0.436  | *** | 0.08 | 0.409  | *** | 0.078 |
| $\ln w_{i,t-1}$                    | 0.226           | *** | 0.070  | 0.243  | ***                 | 0.064 | 0.396  | *** | 0.08 | 0.417  | *** | 0.076 |
| $\ln e_{f,t-1}$                    | 1.963           | *** | 0.282  | 2.038  | ***                 | 0.261 | 2.738  | *** | 0.34 | 3.383  | *** | 0.305 |
| $\ln e_{i,t-1}$                    |                 |     |        | -0.373 | ***                 | 0.088 |        |     |      | -0.192 | **  | 0.091 |
| $FREE_{i,t-1}$                     | 0.012           |     | 0.015  | 0.001  |                     | 0.013 |        |     |      |        |     |       |
| $IRON_{i,t-1}$                     | -0.049          | *** | 0.018  | -0.024 |                     | 0.023 |        |     |      |        |     |       |
| $DICT_{i,t-1}$                     | 0.107           | *** | 0.019  | 0.121  | ***                 | 0.018 | -0.044 | **  | 0.02 | 0.027  |     | 0.034 |
| $WAR_{i,t-1}$                      | -0.160          | *** | 0.061  | -0.151 | ***                 | 0.053 | 0.191  | *** | 0.04 | 0.027  |     | 0.062 |
| $ASYL_{f,t-1}$                     | -0.041          | *** | 0.011  | -0.019 | *                   | 0.010 | -0.065 | *** | 0.01 | -0.010 |     | 0.015 |
| $\Delta \ln mst_{i,t-1}$           | 0.479           | *** |        | 0.521  |                     | 0.025 | 0.474  | *** | 0.02 | 0.455  | *** | 0.029 |
| $\Delta \ln \frac{w_{ft}}{w_{it}}$ | 0.768           | *** | 0.274  | 1.023  | ***                 | 0.248 | 0.326  |     | 0.32 | 1.748  | *** | 0.256 |
| $\Delta \ln w_{i,t-1}$             | 0.142           |     | 0.274  | 0.197  |                     | 0.247 | 0.248  |     | 0.33 | 1.538  | *** | 0.259 |
| $\Delta \ln e_{f,t-1}$             | 0.784           | **  | 0.332  | 1.181  | ***                 | 0.306 | 1.102  | *** | 0.41 | 1.561  | *** | 0.328 |
| $\Delta \ln e_{i,t-1}$             |                 |     |        | -0.259 |                     | 0.204 |        |     |      | -0.075 |     | 0.156 |
| $\Delta FREE_{i,t-1}$              | 0.010           |     | 0.039  | 0.005  |                     | 0.033 |        |     |      |        |     |       |
| $\Delta IRON_{i,t-1}$              | -0.224          | *** | 0.050  | -0.296 | ***                 | 0.052 |        |     |      |        |     |       |
| $\Delta DICT_{i,t-1}$              | 0.102           | **  | 0.041  | 0.106  | ***                 | 0.035 | -0.034 |     | 0.06 | 0.053  |     | 0.049 |
| $\Delta WAR_{i,t-1}$               | -0.077          |     | 0.069  | -0.088 |                     | 0.059 | 0.123  | *** | 0.04 | 0.037  |     | 0.042 |
| $\Delta ASYL_{f,t-1}$              | 0.018           |     | 0.020  | 0.034  | **                  | 0.017 | 0.016  |     | 0.02 | 0.078  | *** | 0.017 |
| obs.                               | 984             |     |        | 869    |                     |       | 1107   |     |      | 450    |     |       |
| groups                             | 24              |     |        | 24     |                     |       | 27     |     |      | 18     |     |       |
| balanced                           |                 |     |        |        |                     |       |        |     |      |        |     |       |
| panel                              | yes             |     |        | no     |                     |       | yes    |     |      | no     |     |       |
| $R^2$                              | 0.46            |     |        | 0.51   |                     |       | 0.44   |     |      | 0.58   |     |       |

Notes: \*\*\*, \*\*, \* denote the 1%-, 5%- and 10% significance levels, respectively.– The dependent variable is the first difference of the log of the migration stock,  $d \ln mst_{it}$ , normalized by the home country population.– All regressions are estimated by weighted OLS, using the log of the population size as weights.– All Regressions include a dummy variable which control for a statistical break in the time series in 1987.– The  $F(23,941)$  statistic testing for country-specific fixed effects is 3.33\*\*\* in the first, the  $F(23,824)$  statistic is 4.26\*\*\* in the second, the  $F(26,1065)$  statistics is 3.88\*\*\* in the third, and the  $F(17,415)$  statistic is 8.03\*\*\*, which rejects the  $H_0$  of no) country-specific fixed effects at the 1%-level in all regressions.

Table 5: List of variables

---



---

|                         |  |
|-------------------------|--|
| $mst_{it}$              | Foreign population from country $i$ in Germany as a percentage of sending country population. Source: Statistisches Bundesamt (2010a).   |
| $m_{it}$                | Net migration flow from country $i$ into Germany as a percentage of sending country population. Source: Statistisches Bundesamt (2010b).   |
| $im_{it}$               | Gross migration inflow from country $i$ into Germany as a percentage of sending country population. Source: Statistisches Bundesamt (2010b).                                       |
| $out_{it}$              | Gross migration outflow from country $i$ into Germany as a percentage of sending country population. Source: Statistisches Bundesamt (2010b).                                      |
| $\frac{w_{ft}}{w_{it}}$ | Ratio of the GDP per worker of Germany to the GDP per worker of sending country $i$ measured in PPP. Sources:<br>Heston, Summers, and Aten (2009), World Bank (2010), OECD (2010). |
| $w_{it}$                | GDP per worker of sending country $i$ measured in PPP. Sources:<br>Heston, Summers, and Aten (2009), World Bank (2010), OECD (2010).   |
| $e_{ft}$                | Employment rate of Germany in % (100 minus the unemployment rate in %).<br>OECD (2010), World Bank (2010).   |
| $e_{it}$                | Employment rate of sending country $i$ in % (100 minus the unemployment rate in %).<br>OECD (2010), World Bank (2010).   |
| $FREE_{it}$             | Dummy variable for participation of country $i$ in the free movement of the European Common Market. Source: Own compilation.   |
| $IRON_{it}$             | Dummy variable for emigration barriers in country $i$ during the 'Iron Curtain' period. Source: Own compilation.   |
| $DICT_{it}$             | Dummy variable for dictatorship in country $i$ . Source: Own compilation based on FreedomHouse (2008).   |
| $WAR_{it}$              | Dummy variable for war in country $i$ . Source: Own compilation.   |
| $ASYL_{ft}$             | Dummy variable for the asylum restrictions introduced in Germany from 1992 onwards. Source: Own compilation.   |

---



---

Table 6: Descriptive Statistics

| Variable                    | Obs  | Mean   | Std. Dev. | Min     | Max    |
|-----------------------------|------|--------|-----------|---------|--------|
| $mst_{it}$                  | 2193 | 0.428  | 0.779     | 0.000   | 4.565  |
| $\ln mst_{it}$              | 2193 | -2.329 | 1.990     | -9.915  | 1.518  |
| $m_{it}$                    | 2193 | 0.012  | 0.071     | -0.611  | 1.122  |
| $im_{it}$                   | 2193 | 0.056  | 0.114     | 0.000   | 1.696  |
| $\ln im_{it}$               | 2193 | -4.221 | 1.868     | -12.946 | 0.528  |
| $out_{it}$                  | 2193 | 0.044  | 0.081     | 0.000   | 0.847  |
| $\ln out_{it}$              | 2193 | -4.555 | 2.000     | -13.445 | -0.166 |
| $\frac{w_{ft}}{w_{it}}$     | 2193 | 4.777  | 6.225     | 0.541   | 49.249 |
| $\ln \frac{w_{ft}}{w_{it}}$ | 2193 | 1.022  | 0.974     | -0.615  | 3.897  |
| $w_{it}$                    | 2193 | 20440  | 15268     | 968     | 64918  |
| $\ln w_{it}$                | 2193 | 9.542  | 0.986     | 6.876   | 11.081 |
| $e_{ft}$                    | 51   | 94.02  | 3.02      | 88.85   | 99.43  |
| $\ln e_{ft}$                | 51   | 4.543  | 0.032     | 4.487   | 4.599  |
| $e_{it}$                    | 1379 | 92.77  | 5.39      | 69.00   | 99.94  |
| $\ln e_{it}$                | 1379 | 4.528  | 0.061     | 4.234   | 4.605  |