

# **Selective in-migration and income convergence and divergence across Belgian municipalities**

Ludo Peeters\*

\* KIZOK – Research Centre for Entrepreneurship and Innovation  
Hasselt University  
B-3590 Diepenbeek (Belgium)  
Tel. ++32 11 26 86 35  
Fax ++32 11 26 87 00  
E-mail: [ludo.peeters@uhasselt.be](mailto:ludo.peeters@uhasselt.be)

Acknowledgements – We thank Yves Surry, Marc Tiri, the editor, and two anonymous referees for their valuable comments and suggestions. Any errors or omissions remain our responsibility.

## ABSTRACT

In this paper, we present an empirical study of the relationship between aggregate migration and the income distribution across locations in a region. We use cross-sectional data on growth of per-capita taxable incomes, over the period 1991-2000, from the municipalities of the Belgian province of Limburg. We use a non-standard  $\beta$ -convergence framework, built on the notion of ‘weak absolute convergence’, to test the hypothesis of selective income migration. The model we develop allows us to separate out observed and unobserved heterogeneity in convergence rates. Observed heterogeneity is incorporated by including net in-migration and initial-education variables. Unobserved heterogeneity is related to amenity differentials across locations. The model is estimated by using the generalized maximum entropy method, which allows for the ‘non-parametric’ estimation of the effects of the unobserved factors. While we find an average  $\beta$ -convergence rate of about 4.2% annually, the heterogeneity in  $\beta$ -convergence rates is consistent with an increasing dispersion ( $\sigma$ -divergence) of income levels, due to selective in-migration. Our empirical evidence indicates that, overall, the initially high-income municipalities (including the main regional urban centres) as well as their immediate neighbours gained the most from selective income migration. We conclude that patterns of in-migration in the 1990s provided the basis for a cumulative process of divergence of per-capita incomes and a growing spatial concentration of high-income locations in the amenity-rich, south-western part of the province.

Key words: Taxable incomes Income migration Amenities Heterogeneity  
Weak absolute convergence Generalized maximum entropy

JEL classifications: C13, C14, R12, R23

## INTRODUCTION

The relationship between migration and income growth at the regional level has gained much attention in the literature. Most of the research on internal migration has focused primarily on the *determinants* of migration. A large body of literature has debated the relative importance of economic/job opportunities versus local amenities as determinants of migration, and ample empirical evidence seems to support both perspectives (e.g., PORELL, 1982; MUESER and GRAVES, 1995). In contrast, relatively little has been published in the regional-science literature regarding the *consequences* of migration, in terms of, for example, per-capita incomes in the destination areas (e.g., GREENWOOD, 1985; PLANE and BITTER, 1997; ARONSSON *et al.*, 2001; CUSHING and POOT, 2004). In the present paper, we present an empirical study of the effect of population movements on the re-distribution of per-capita disposable incomes in a region, and investigate whether there is some linkage with the spatial distribution of local amenities and other quality-of-life factors (e.g., MARGO, 1992; CROMARTIE and NORD, 1997; PLANE, 1999; SHUMWAY and OTTERSTROM, 2001).<sup>1</sup> The analysis is based on cross-sectional data from the 44 municipalities of the province of Limburg, located in the north-eastern part of Belgium.

Migration represents more than a simple re-distribution of (homogeneous) people; it also affects the *composition* of the population both in origins and destinations. Although migrants are economically and demographically diverse, using aggregate migration data usually prevents one from sorting out different categories of in- and out-migrants, and assessing their relative income position, work status, education, age, and so on (e.g., FRIEDBERG and HUNT, 1995; PLANE and HEINS, 2003). In this paper, we are primarily concerned with the income stratification of aggregate migration and the way in which the selectivity of who moves where affects the income distribution across locations.<sup>2</sup> The migratory process can change the relative incomes across locations in two ways. Firstly, if more people move into a particular location than leave it, and assuming the average incomes of both in- and out-migrants are equal (but higher than the average incomes of non-migrants), then there will be a net (positive) transfer of income: the per-capita income at this particular location tends to rise. Secondly, the per-capita incomes of in- and out-migrants may be different. If in-migrants have higher per-capita incomes than out-migrants, there will also be a net (positive) transfer

of income. In other words, per-capita incomes may be affected by net in-migration through a quantity effect and/or a composition effect (e.g., SHIOJI, 2001).

Without knowledge of the incomes of migrants, however, it is not possible to identify the distinctive spatial patterns of migration for high-income versus low-income people. But, if location preferences and migration paths, in terms of origins/destinations, consistently vary across income groups of migrants, migration may lead towards a spatial re-distribution of per-capita incomes in the region. In the present paper, we hypothesize, as CROMARTIE and NORD (1997), among others, do, that migrants moving into amenity-rich areas have higher per-capita incomes than out-migrants (and possibly also non-migrants), and, therefore, these net-receiving areas may gain income from both the total number (quantity effect) and the income characteristics of in-migrants (composition or differential-income effect). Such a hypothesis is also consistent with the literature on migration and amenities (see, for example, KNAPP and GRAVES, 1989; GOTTLIEB, 1993; MUESER and GRAVES, 1995; BURNLEY *et al.*, 1997; BRUECKNER *et al.*, 1999; MATHUR and STEIN, 2005). In this literature it appears to be a ‘stylized fact’ that local amenities, and other quality-of-life factors, are playing an increasingly important role in migration decisions, and that rising income leads to an increased demand for local amenities (i.e., access to product variety). In an earlier study on internal migration in Belgium (DEVOGELAER, 2004), it was also found that high-income migrants are more sensitive to the supply of amenities rather than, for example, to the presence of job opportunities (see also SHELLEY and KOVEN, 1993; ANJOMANI, 2002; ADAMSON *et al.*, 2004).<sup>3</sup>

If the income stratification of the migratory process is not observable, how can we possibly empirically test whether a selective income-migration has actually taken place? Our research strategy is to provide an *indirect* test by using a  $\beta$ -convergence framework, which will be designed in such a way that it captures the roles of both observed and ‘unobserved’ (by the researcher) heterogeneity shaping the migration paths of people belonging to different income categories. Our argument goes as follows: if high-income migration paths tend to sustain those locations that already had high initial per-capita incomes, then one may expect: (i) divergence of per-capita incomes across municipalities; (ii) growing spatial concentration of high per-capita incomes *in* or *close to* the regional urban centres that are endowed with a strong amenity advantage. Conversely, the spatial growth pattern of per-capita incomes may

reveal some information about the income composition of migration flows to particular destinations.

To the extent that selective migration patterns are related to the spatial distribution of amenities in the region, this may have important long-run implications. Because service amenities are both a cause *and* a consequence of location/migration patterns of different income groups, a process of cumulative causation may take place (where the causation is taken to occur within the context of a particular amenity mix) leading towards further dispersion and spatial concentration of per-capita incomes in the region. In other words, selective income migration may provide the basis for a self-perpetuating process of divergence of per-capita incomes between ‘central’ and ‘peripheral’ areas.

The layout of the paper is as follows. We begin with an exploratory analysis of per-capita incomes across municipalities in the province of Limburg. Next, we present our basic methodology of testing the selective income-migration hypothesis, using a non-standard  $\beta$ -convergence framework. Then, we specify the empirical model. Finally, we present the empirical results. The paper concludes with a summary and a discussion of some policy issues.

## **EXPLORATORY DATA ANALYSIS**

In this study, we use cross-sectional data on incomes drawn from the official statistics provided by the Administration of Planning and Statistics (APS) of the Flemish government. The data relate to the 44 municipalities of the Belgian province of Limburg, for the period 1991-2000.<sup>4</sup>

Per-capita taxable income is used as a proxy for the per-capita disposable income (see also, for example, DRENNAN and LOBO, 1999).<sup>5</sup> This income indicator is deflated by the national consumer’s price index, expressed in constant 1996 prices.<sup>6</sup> Taxable incomes are recorded and reported on a place-of-residence basis, and, hence, represent the relative prosperity of the resident populations in each municipality (including the local residents who commute to their workplace in other municipalities).

[Table 1 about here]

Table 1 reports some summary statistics on the distribution of per-capita incomes (in 1,000 EURs), at various points of time over the full period 1991-2000. It can be observed that the dispersion of per-capita incomes across municipalities, as measured by the standard deviation or the coefficient of variation, has increased substantially in the 1990s, with a noticeable jump in 2000. Obviously, the study period has been one of  $\sigma$ -divergence. To assess the evolution of the shape of the distribution, Table 1 also presents measures of skewness and kurtosis. The distribution shows a trend of declining positive skew, which means that those municipalities with the lowest initial per-capita incomes are increasingly lagging behind. Also, the kurtosis of the distribution decreases considerably, which means that the distribution becomes flatter and more spread out. Apart from examining the dynamics of income dispersion, it is also instructive to examine the extent of the intra-distributional mobility over time by looking at the changes in the ranking of the municipalities with respect to per-capita income. A suitable measure of intra-distributional mobility is based on Kendall's  $W$  index of rank concordance, which is shown to be closely related to the notion of  $\beta$ -convergence (BOYLE and MCCARTHY, 1997). The figures in Table 1 reveal that there was no great movement within the distribution of per-capita incomes across municipalities. In fact, we had to reject any significant adjustment in the ordinal rankings at the 1% level. Also, the Spearman's rank correlation ( $\rho$ ) between the per-capita income levels in 1991 and 2000 is equal to 0.760 ( $p < 0.001$ ). Finally, an increasing spatial clustering of low- and high-income municipalities can be discerned by looking at the Moran scatter plots in Fig. 1. Specifically, the Moran's  $I$  increased from 0.029 ( $p = 0.594$ ) in 1991 (no spatial association) to 0.370 ( $p < 0.001$ ) in 2000. While Hasselt (the capital of the province) occupied a strong lead position in 1991, its neighbouring municipalities seem to be catching up fairly rapidly.

[Fig. 1 about here]

[Fig. 2 about here]

In order to provide a better visual representation, the spatial distribution of the initial per-capita incomes (1991) and the average annual growth rates of per-capita incomes (1991-2000) are portrayed in Fig. 2. Each municipality is shaded according to the quartile it belongs to. As

can be seen from Panel A of Fig. 2, the highest initial per-capita incomes are found in the main regional urban centres Hasselt (the capital of the province), Sint-Truiden, and Tongeren, which are known to be endowed with strong amenity advantages (see, for example, the urban typology developed by VAN HECKE, 1998).<sup>7</sup> Besides, high per-capita incomes are also found in some of the municipalities adjacent to Hasselt and Genk, while a number middle-high per-capita income levels show up in the northern part of the province. Conversely, the growth rates of per-capita incomes over the period 1991-2000 are more spatially clustered; the corresponding Moran's  $I$  is equal to 0.332 ( $p < 0.001$ ). The most striking feature in Panel B of Fig. 2 is the distinctive pattern of black and dark grey shading in the south-west of the province, which indicates a spatial clustering of fast growth rates *in or close to* the main regional urban centres. We hypothesize that an important reason of the superior growth performance of this cluster is that the constituent municipalities have some favourable endowments, in terms of supply of (or proximity to) urban-based service amenities and other quality-of-life factors, making them attractive for high-income in-migrants and non-migrants.<sup>8</sup>

## BASIC METHODOLOGY

As said in the introductory section, the main purpose of this paper is to provide an indirect test of the selective income-migration hypothesis using a 'non-standard'  $\beta$ -convergence framework. In contrast with the mainstream convergence literature (e.g., BARRO and SALA-I-MARTIN, 2003; ISLAM, 2003), this study looks at growth and convergence from a cumulative-causation, rather than a neo-classical, perspective (see also, for example, ROBERTS, 2004). In the cumulative-causation perspective, regional or local income-growth disparities may persist, or even increase, due to positive feedback effects, which, then, could explain why we observe  $\sigma$ -divergence of per-capita incomes across the municipalities in the province of Limburg (see previous section, Table 1).

Most studies on  $\beta$ -convergence involve growth regressions based on the notion of either absolute (unconditional) or conditional convergence. In estimating such growth regressions, researchers explicitly or implicitly impose the assumption that all economies are approaching to a common steady-state growth path or to their own steady-state growth paths, and do so at the *same* rate. As a result, it is thought meaningful to estimate a single rate of convergence.

However, as shown by EVANS (1997), among others, the conventional approach can produce consistent estimates only under highly restrictive conditions. One such condition is that the control variables in the growth regression should account for all cross-sectional heterogeneity. Given that such a condition is usually not satisfied in practice, adopting a single linear growth model may lead to severely biased estimates and incorrect inferences. Moreover, the use of a cross-sectional framework raises a problem concerning the correct interpretation of the conditional-convergence rate. If there is no common equilibrium growth rate, then the whole notion of conditional  $\beta$ -convergence has little economic meaning and, hence, the ‘across interpretation’ of  $\beta$  is no longer valid (see ISLAM, 2003, p. 321). A negative sign for  $\beta$  is merely taken as evidence showing that, *other things held constant*, units starting with lower levels of income grow faster. Yet, this does not say anything about convergence as such.

### *Weak Absolute Convergence*

Given the aforementioned problems, the  $\beta$ -convergence model we use is built on the (new) notion of ‘weak absolute convergence’ (WAC). Under WAC, it is assumed that the cross-sectional units converge to the *same* balanced-growth path, but at *different speeds*.<sup>9</sup> In other words, WAC implies that the relationship between initial income and subsequent income growth is varying (heterogeneous) over the municipalities.<sup>10</sup>

Formally, the WAC model we propose is as follows:

$$g_i = \alpha + \beta_i y_{i0} = \alpha + (\gamma' \mathbf{z}_i + \delta_i) y_{i0}, \quad i = 1, \dots, N \quad (1)$$

where the index  $i$  designates the unit of analysis (municipality),  $g_i$  is the average annual growth of per-capita income between the base year 0 and the terminal year  $T$ , calculated as the natural (log) growth rate (that is,  $g_i = \ln(Y_{iT}/Y_{i0})/T$ ),  $y_{i0}$  is the natural log of the per-capita income level at base year 0 (that is,  $y_{i0} = \ln(Y_{i0})$ ),  $\mathbf{z}_i$  is a vector of location-specific variables that control for the *observed* cross-sectional heterogeneity, and the  $\delta_i$ s are location-specific terms (with zero mean and finite variance), which account for the *unobserved* heterogeneity. The WAC specification means that each municipality has its own initial-income coefficient,  $\beta_i = \gamma' \mathbf{z}_i + \delta_i$ , and the corresponding annual rates of convergence are



calculated as  $r_i = -\ln(1 + \beta_i T)/T$ . It is a well-known result that convergence requires that the  $\beta$  coefficient is negative, and convergence is faster the more negative  $\beta$  is.

Before moving on, some remarks on equation (1) are in order. Firstly, the variables in  $\mathbf{z}_i$  are not to be considered as ‘explanatory variables’ in the usual sense; instead of shifting the growth regression, these covariates help to explain the variation in the  $\beta$  coefficients (thereby rotating, so to speak, the growth regression), and hence, only affect the growth of per-capita incomes indirectly. Secondly, we eliminate the ‘true’ error term, which is supposed to capture the effect of omitted variables that are *not* correlated with initial income, and assume that all unobserved effects are picked up by the location-specific  $\delta_i$ s (which is just an alternative way of injecting heterogeneity across observational units). Finally, the varying  $\beta$  coefficients – and, by implication, also the unobserved effects – are allowed to be arbitrarily correlated (to interact) with the initial-income variable.<sup>11</sup>

### *Unobserved effects*

The WAC model in equation (1) is designed to separate out the role of unobserved effects, by including the separate location-specific terms  $\delta_i$ . These are supposed to capture the effects on per-capita income growth from a multitude of unobserved factors which may affect the relative attractiveness (or repulsiveness) of a location for different income groups of both in- and non-migrants to reside, and which may particularly be related to local service amenities and other quality-of-life factors.

Whilst in the migration literature it is common practice to confine these factors to a single attribute or an *ad-hoc* list of selected attributes, mostly related to climate and other environmental factors (see, for example, DELLER *et al.*, 2001; ARONSSON *et al.*, 2001), we prefer to (implicitly) include unit-specific dummies in the model to capture unobserved local factors, and impose some spatial structure on their effects, as will be explained shortly. After all, climate evidently does not play a role in the present context, whilst proximity to service amenities and other quality-of-life factors (e.g., landscape features, quietness, etc.), which usually present themselves as a mix or package (of possibly interacting attributes), are extremely difficult to quantify and thus hard to control for.

### *Spatial interaction (spill-over) effects*

Given that our units of analysis are municipalities, the explicit recognition of spatial interaction or spill-over effects is considered to be critically important in estimating the WAC model. Such cross-border effects may be caused by the hypothesized spatial correlation among the unobserved characteristics of neighbouring municipalities.

Assuming spatial dependence among the unobserved effects,  $\delta_i$ , we write

$$\delta_i = \lambda \sum_{j=1}^N \tilde{w}_{ij} \delta_j + u_i, \quad i = 1, \dots, N \quad (2)$$

where  $\lambda$  is the coefficient of spatial correlation. The spatial weights  $\tilde{w}_{ij}$  are defined in two different ways. In the first case, we assume that spatial dependence is at work only among immediately neighbouring municipalities; that is,  $\tilde{w}_{ij}$  is the  $(i,j)$ -th element of a row-standardized contiguity matrix. In the original (binary) contiguity matrix,  $w_{ij} = 1$  if municipalities  $i$  and  $j$  are contiguous, and  $w_{ij} = 0$  otherwise. In contrast with standard practice, we also set all diagonal elements equal to one in the original contiguity matrix ( $w_{ii} = 1$ ). In the second case, we assume a more general form of spatial dependence, where the interaction varies with the inverse of the physical distance between the municipalities; that is,  $\tilde{w}_{ij}$  is the  $(i,j)$ -th element of an inverse-distance matrix, where  $\tilde{w}_{ii} = 1/d_{ii}$ ,  $d_{ii}$  is the ‘intra-municipal distance’ of municipality  $i$ ,<sup>12</sup>  $\tilde{w}_{ij} = 1/d_{ij}$ , and  $d_{ij}$  is the Euclidean distance (in kilometres) between the centres of the municipalities  $i$  and  $j$ . Note that by incorporating equation (2) into equation (1), we (re-)introduce a noise component  $u_i$ , which is assumed to have zero mean and finite variance.

The term associated with  $\lambda$  in equation (2), which can be written as  $c_i = \sum_j \tilde{w}_{ij} \delta_j$ , represents some sort of ‘centrality index’ of attractiveness (or repulsiveness) associated with ‘clusters’ of locations around municipality  $i$ . This index  $c_i$  is defined in such a way that the unobserved effect associated with a particular municipality  $i$  is the average of its ‘own’ unobserved effect and the unobserved effects of its adjacent municipalities (in the case of

contiguity), or a linear combination of the unobserved effects from all other municipalities (in the case of distance), where the weights attached to each municipality  $j$  decrease with distance to municipality  $i$ . The centrality index  $c_i$  may thus capture the ‘competition’ that each location faces from other locations due to unobserved factors related to amenity differentials.<sup>13</sup> Specifically, if  $\lambda > 0$ , then *agglomeration* forces are at work, which means that locations in close proximity to other attractive locations are also attractive (centrality); if  $\lambda < 0$ , then *competition* effects are present, implying that locations in close proximity to attractive locations are less attractive (de-centrality).

Given the definition of the unobserved effects, the growth regression in equation (1) can be written as

$$g_i = \alpha + (\gamma' \mathbf{z}_i + \lambda \sum_{j=1}^N \tilde{w}_{ij} \delta_j) y_{i0} + u_i y_{i0}, \quad i = 1, \dots, N \quad (3)$$

Next, observing from equation (1) that  $\delta_j = (g_j - \alpha - (\gamma' \mathbf{z}_j) y_{j0}) / y_{j0}$ , equation (3) can be reformulated as

$$g_i = \alpha + (\gamma' \mathbf{z}_i) y_{i0} + \lambda \sum_{j=1}^N \tilde{w}_{ij} \left[ \frac{g_j - \alpha - (\gamma' \mathbf{z}_j) y_{j0}}{y_{j0}} \right] y_{i0} + u_i y_{i0}, \quad i = 1, \dots, N \quad (4)$$

As a result, the location-specific unobserved effects have been ‘factored out’ in equation (4), and are now fully defined in terms of the other parameters  $\alpha$ ,  $\gamma$ , and  $\lambda$ . The estimated  $\hat{\delta}_i$ s can subsequently be obtained as

$$\hat{\delta}_i = \hat{\lambda} \sum_{j=1}^N \tilde{w}_{ij} \left[ \frac{g_j - \hat{\alpha} - (\hat{\gamma}' \mathbf{z}_j) y_{j0}}{y_{j0}} \right], \quad i = 1, \dots, N \quad (5)$$

We now move on to presenting the empirical specification of the WAC model.

## EMPIRICAL MODEL SPECIFICATION

In any actual empirical application the choice of appropriate variables is a difficult issue, since economic theory is not always informative about which variables belong to the ‘true regression’, and many variables that are specific to the units of analysis may simply not be observable. While acknowledging this problem, we adopt a parsimonious specification of the growth regression, rather than ‘trying’ a wide set of possible variables, and control for the unobserved heterogeneity through exploiting the flexible design of the WAC model.<sup>14</sup>

As our primary focus is on the effect of migration on the growth of per-capita incomes at the municipal level as well as on the associated rates of convergence, we include the average annual net in-migration rates (per 100 inhabitants), over the period 1991-2000 (data drawn from APS).<sup>15</sup> Besides migration, we also include the initial (1991) level of education – evidently an important determinant of income – as another important variable to account for observed heterogeneity. The proxy used for initial educational attainment is the percentage of the (resident) working population in a municipality with a post-secondary or university degree (data drawn from the 1991 National Census of Population). The descriptive statistics given in Table 2 show a wide variation in both net in-migration rates and initial levels of educational attainment across municipalities.

It is clear, though, that net in-migration and education can account for only part of the variation in growth and convergence rates. However, by putting forward a flexible varying-coefficients model that also accounts for unobserved effects (through the  $\delta_i$ s), we are able, at least to some extent, to alleviate any omitted-variable bias.

[Table 2 about here]

Thus, the final specification of the empirical WAC model is

$$g_i = \alpha + (\gamma_0 + \gamma_1 m_i + \gamma_2 x_{i0}^{-1} + \lambda \sum_{j=1}^N \tilde{w}_{ij} \delta_j) y_{i0} + u_i y_{i0}, \quad i = 1, \dots, N \quad (6)$$

where  $m_i$  is the average annual net in-migration rate in location  $i$ , and  $x_{i0}$  is the initial level of education in location  $i$ . Note, in passing, that by entering the educational-attainment variable

*inversely* in the regression, we allow for diminishing marginal (social) ‘returns’ of education in terms of subsequent growth performance (see also ISLAM, 2003, footnote 41, p. 347).

If all  $\beta_i$ s are negative, then there is overall convergence to the common steady-state growth path of per-capita incomes, although possibly at different speeds. The sign of  $\hat{\gamma}_1$  is generally not known *a priori*. If  $\hat{\gamma}_1 > 0$ , then net in-migration has a positive impact on the growth of per-capita income, while at the same time slowing down convergence. If such is the case, this would be in support of the selective income-migration hypothesis. Conversely, if  $\hat{\gamma}_1 < 0$ , then net in-migration has a negative impact on growth, and contributes to convergence. On the other hand, we expect to find a negative value for  $\hat{\gamma}_2$ , hence  $\partial g_i / \partial x_{i0} = -\hat{\gamma}_2 / x_{i0}^2 > 0$ , indicating a positive impact of initial education on subsequent growth of per-capita income, along with a ‘dampening’ effect on convergence. Furthermore, we expect to find  $\hat{\delta}_i > 0$  for those municipalities with a relative amenity advantage, and  $\hat{\delta}_i < 0$  for those with a relative amenity disadvantage. Finally, given that amenity-related unobserved effects may possibly be at work (spill-over) beyond the municipal borders, we expect  $\lambda$  to be positive, since many of the unobserved characteristics are likely to be positively correlated over space, given our detailed data set. For example, if a municipality is characterized by the presence of high-quality schools, its neighbouring municipalities will also be close to these high-quality schools. Or, if a municipality is located in open space or rolling hills, adjacent municipalities are likely to be also.

### *Generalized Maximum Entropy method*

The empirical model will be estimated by applying the Generalized Maximum Entropy (GME) method (GOLAN *et al.*, 1996). Whilst sidestepping the basics of entropy econometrics, the GME formulation of our empirical WAC model is outlined in the Appendix to this paper (for a concise summary on maximum-entropy estimation, see also FRASER, 2000).

Our principal motivation for applying GME is that this method allows us to control for unobserved or individual fixed effects within a cross-sectional setting, permitting ‘non-parametric’ shifts of the  $\beta$  coefficient. In other words, the unobserved effects are allowed to vary ‘freely’, rather than in accordance with some distribution function (recall that the  $\delta_i$ s in

equation (1) can be viewed as the coefficients of observation-specific dummies). Evidently, estimating the individual  $\delta_i$ s would have been ruled out when using classical estimation methods, due to the ill-posed or underdetermined nature of the model (given that the number of unknown coefficients is larger than the number of observations). In contrast, GME is known to be ‘immune’ to this dimensionality problem, and thus allows us to uniquely identify the unobserved effect for each individual municipality in a cross-sectional setting (rather than merely the mean and the variance of an assumed distribution, as would be the case in a classical random-coefficients model).<sup>16</sup>

## EMPIRICAL RESULTS

This section contains the results from the GME estimation of the empirical WAC model.<sup>17</sup> In Table 3 we report the GME point estimates of the coefficients, along with their bootstrapped standard errors and 90% confidence intervals, for both the contiguity-based and the distance-based weighting alternatives (Panel A and Panel B, respectively).<sup>18</sup> Obviously, the estimation results are affected by the type of spatial weighting adopted. However, given the better model fit obtained from using distance-based spatial weighting (the pseudo- $R^2$ , defined as the square of the correlation between the actual and the predicted growth rates, is equal to 0.54, which is quite high in a cross-sectional setting), we limit ourselves hereafter to discussing the results reported in Panel B of Table 3.<sup>19</sup>

The results indicate that all the estimated coefficients are significant (at the 10% level), and that the data are consistent with the WAC hypothesis:<sup>20</sup> all  $\hat{\beta}_i$ s are negative, with mean equal to -0.0348 and standard deviation equal to 0.0025. The variance of the  $\beta$  coefficients can be decomposed by using the expression:  $\sigma_{\beta}^2 = \hat{\gamma}_1^2 \sigma_m^2 + \hat{\gamma}_2^2 \sigma_{x_{t-1}}^2 + \sigma_{\delta}^2 + \text{Cov}'s$ . Based on this decomposition, we find that 48.1% of the variation in the  $\beta$  coefficients can be attributed to the variation in net in-migration (6.2%) and initial education (41.9%), whereas 42.7% is due to the unobserved effects. The covariances account for the remaining part of the variance (9.2%).

Based on the GME estimates, the annual convergence rates across municipalities are in the range of 3.6% (minimum) to 5.1% (maximum), with a mean value of 4.2%. This means that there is overall convergence to a common growth path, although at different speeds.

[Table 3 about here]

Furthermore, the estimated coefficients associated with the net in-migration and initial-education variables have the expected signs. Both conditioning variables have a positive effect on the growth rate of per-capita income (recall that initial education is included inversely in the model):  $\partial g_i / \partial m_i = \hat{\gamma}_1 y_{i0} > 0$  and  $\partial g_i / \partial x_{i0} = -\hat{\gamma}_2 y_{i0} / x_{i0} > 0$ , given that  $\hat{\gamma}_1 > 0$  (0.2537) and  $\hat{\gamma}_2 < 0$  (-0.0020), respectively. What is particularly interesting about these results is that the positive effect of net in-migration on per-capita income growth *increases* with the level of initial income ( $\partial^2 g_i / \partial m_i \partial y_{i0} = \hat{\gamma}_1 > 0$ ), *ceteris paribus*. In other words, initially ‘rich’ municipalities tend to benefit most from selective in-migration. Also, the positive effect of initial education increases with the level of initial income ( $\partial^2 g_i / \partial x_{i0} \partial y_{i0} = -\hat{\gamma}_2 / x_{i0} > 0$ ), whereas the gain in per-capita income *decreases* (negative exponentially) with its own initial level ( $\partial^2 g_i / \partial x_{i0}^2 = \hat{\gamma}_2 y_{i0} / x_{i0}^2 < 0$ ), *ceteris paribus*. The latter may be interpreted as indicating (i) diminishing marginal (social) returns to education in terms of subsequent income growth, (ii) a ‘waste’ of highly-educated people who decided to move to other municipalities during the study period, or (iii) a combination of the two.

While net in-migration and initial education appear to be important drivers of per-capita income growth, the results also imply that both variables tend to *reduce* the convergence speed to the common balanced-growth path, as  $\beta$  becomes less negative:  $\partial \beta_i / \partial m_i = \hat{\gamma}_1 > 0$  and  $\partial \beta_i / \partial x_{i0} = -\hat{\gamma}_2 / x_{i0}^2 > 0$ . In other words, higher values for net in-migration and initial education imply that the growth rate of per-capita income continues to be larger than the balanced growth path for quite an extended period of time, whilst lower values imply that the growth rate of per-capita income converges more rapidly to the balanced-growth path (the correlation between convergence rates and growth rates is equal to -0.775).

### *Unobserved effects*

The unobserved effects, which are supposed to be captured by the  $\hat{\delta}_i$ s, appear to play an important role in accounting for the variation in growth rates of per-capita income. The values of the  $\hat{\delta}_i$ s range between -0.0027 and 0.0031 (to save space, we do not report all the individual  $\hat{\delta}_i$ s here). A positive value of  $\hat{\delta}_i$  is associated with a higher income-growth rate (and a lower convergence speed), given the levels of net in-migration and initial education, and *vice versa*. Interestingly, the unobserved effects are not randomly distributed over space. Consistent with our prior expectations, the estimated autocorrelation coefficient  $\hat{\lambda}$  is positive and equal to 0.656, while the Moran's  $I$  for the individual  $\hat{\delta}_i$ s is equal to 0.806 ( $p < 0.001$ ).<sup>21</sup> As a result, there appear to be strong agglomeration forces at work with respect to the unobserved effects.

A map showing the spatial distribution of the estimated unobserved effects is given in Fig. 3. This map mirrors a strong linkage between the spatial distribution of amenities and the differential rates of income-growth, with a noticeable spatial concentration of highly positive values of  $\hat{\delta}_i$  in the south-western part of the province, in particular along the axis formed by Hasselt (H) and Sint-Truiden (ST), which are the most important trade centres in the province. Another interesting finding is that not only do the regional urban centres exhibit a positive amenity-related, unobserved effect, but also – and even more prominently – the immediately neighbouring municipalities (the highest values of  $\hat{\delta}_i$  are associated with the municipalities marked with black shading and white spot). This finding clearly suggests that (cross-municipality-border) *spill-over* effects are also important for local per-capita income growth (see also, for example, ARONSSON *et al.*, 2001).<sup>22</sup> The latter may be indicative of the fact that at least some high-income migrants (as well as high-income non-migrants) are maximizing their location utility by balancing between the demand for service amenities, largely found in the regional urban centres, and preferences for open space, residential property on large lots, and ‘small-town values’, associated with the lower-density, exurban or rural-urban areas. Thus, in order to solve this apparent trade-off, high-income people may choose to reside in (or swarm out to) exurban or rural-urban areas *that are nonetheless close* to the regional urban



centres – that is, areas on a short commuting distance, so that they can still get the benefits from the amenities supplied in the regional urban centres (see also, for example, BURNLEY *et al.*, 1997; RENKOW and HOOVER, 2000; IRWIN and BOCKSTAEL, 2002; CARRIÓN-FLORES and IRWIN, 2004).<sup>23</sup>

[Fig. 3 about here]

### *Explaining unobserved effects*

How can such differences in unobserved effects across municipalities be explained? We can think of three possible explanations that are particularly relevant in the present context. A first possible explanation is that the values of  $\hat{\delta}_i$ s may partly pick up the unobserved differential-income effect of migration (i.e., composition effect) on local per-capita income growth. Accordingly, the municipalities in the south-west cluster might have been able to attract in-migrants belonging to the highest income stratum, at the expense of those municipalities located in the north-eastern part of the province. Another possible explanation is that a positive value of  $\hat{\delta}_i$  may be indicative of the fact that the corresponding municipality has experienced a shift in local demand conditions, through the influx of high-income migrants, opening up new economic opportunities for non-migrants to improve their income position. Obviously, a growing concentration of high-income people in an area inevitably leads to an increasing demand for (new) local services, such as retailing and other activities servicing the local population. This holds particularly for those municipalities in the exurban and rural-urban fringe areas, which entered the 1990s with relatively low initial per-capita income and low amenity endowments, but which are located at close proximity to the initially amenity-rich regional urban centres. Finally, the inflow of high-income people affects the ‘education/skill content’ of the local population or labour force.<sup>24</sup> Given an increasing share of advanced-degree-holding people in a municipality, one can expect that more people have better prospects for improving their personal income position over the study period, which may, in turn, give rise to a higher growth rate of per-capita income.

In sum, it appears as though the constituent municipalities of the south-western spatial cluster benefited the most from amenity-induced, selective in-migration. This, in turn, may

have ‘paved the way’ for these municipalities, particularly those adjacent to the main regional urban centres, to appreciably strengthen their economic base over the study period. Accordingly, the estimated  $\delta_i$ s do more than just reflecting the effects of fixed amenity endowments; they may as well be indicative of some of the dynamics, including changes in local amenity endowments, that might have taken place throughout the study period, and which could explain the superior growth performances of those municipalities adjacent to the amenity-rich regional urban centres.

### *WAC and $\sigma$ -divergence of per-capita income levels*

Even though we find evidence of overall convergence to a common equilibrium *growth* path, the heterogeneity in convergence rates appears to be a consistent with an increasing cross-sectional dispersion ( $\sigma$ -divergence) of income *levels* – provided, of course, that the observed pattern of the 1990s continues to prevail.

Given our empirical estimates from the WAC model, we can easily predict the limiting distribution of the steady-state incomes. Based on CANOVA and MARCET (1995), we can re-write equation (1) as

$$\ln(Y_{iT}/Y_{i0})/T = \alpha - ((1 - e^{-r_i T})/T) \ln Y_{i0} \quad (7)$$

where  $Y_{it}$  is the income level in year  $t$  for municipality  $i$ , and  $\beta_i = -(1 - e^{-r_i T})/T$ . After some rearrangements, we get:  $\ln Y_{iT} = \alpha T + (1 + \beta_i T) \ln Y_{i0}$ . Then, by using discrete time notation and setting  $\rho_i = 1 + \beta_i$  ( $0 < \rho_i < 1$ ), the above equation can be written as:  $\ln Y_{iT} = \alpha + \rho_i \ln Y_{i,T-1} = \alpha + \rho_i(\alpha + \rho_i \ln Y_{i,T-2}) = \alpha + \alpha \rho_i + \rho_i^2(\alpha + \rho_i \ln Y_{i,T-3})$ , etc., which means that  $\ln Y_{iT}$  converges monotonically to  $\ln Y_i^* = \alpha/(1 - \rho_i) = -\alpha/\beta_i$  as  $T$  becomes large, and, by taking the anti-logs,  $Y_{iT}$  converges to

$$Y_i^* = \exp(-\alpha/\beta_i) \quad (8)$$

as  $T$  becomes large.<sup>25</sup> Hence, the lower the convergence speed (i.e., the less negative  $\beta_i$  is), the higher will be the corresponding steady-state income, and *vice versa*. Based on our results,

the mean of the predicted steady-state incomes is equal to 18.0, with a standard deviation of 3.7 (implying a coefficient of variation of 20.3%), and a minimum and maximum value of 11.5 and 25.1, respectively.

To assess the relative impacts of net in-migration and initial education on the steady-state per-capita income levels, we calculate ‘quasi-elasticities’, defined as the percentage change in  $\hat{Y}_i^*$  due to a 1% point change in net in-migration and initial education, evaluated at the sample means ( $\bar{m} = 0.0028$ ,  $\bar{x}_0 = 0.2109$ , and given  $\bar{\delta} = 0$ ). If the net in-migration rate increases by 1% point, the predicted steady-state income level will increase, on average, by about 26.0%, *ceteris paribus*. Note, however, that an increase in the net in-migration rate of 1% point is quite large, given that the mean value of the average annual net in-migration rates over all municipalities is only about 0.28%. Hence, an increase of the net in-migration rate by, say, 0.25% points (i.e., almost doubling the average net in-migration rate), gives rise to a 6.5% increase, on average, of the predicted steady-state income level. On the other hand, if initial education increases by 1% point, the predicted steady-state income level increases, on average, by 3.4%, *ceteris paribus*. Also, the estimated unobserved effects are quite substantial (albeit, at first sight, the  $\hat{\delta}_i$ s may seem to be small in magnitude). Evaluated at the sample means of net in-migration and education, a value of  $\hat{\delta}_i$  equal to, say, 0.001 (-0.001) (the cut-off points used in Fig. 3) is associated with an increase (decrease) of the predicted steady-state income level by 8.8% (7.7%).

Fig. 4 shows the Moran scatter plot of the steady-state income levels predicted by our estimated WAC model, which can be compared with the plots shown in Fig. 1. The Moran’s  $I$  for the steady-state incomes takes a high value of 0.596 ( $p < 0.001$ ), which indicates a growing (global) spatial concentration of high versus low per-capita levels. From Fig. 4 it can also be seen that, for example, Hasselt is gradually losing its strong lead position to the benefit of its immediately surrounding municipalities (Herk-de-Stad in the west, Alken in the south, and, to a somewhat lesser extent, Zonhoven in the north). This finding possibly indicates that short-distance residential mobility (or swarming) of high-income people from the regional urban centres, such as Hasselt, to the exurban and rural-urban fringe areas may also have been important.

[Fig. 4 about here]

Finally, Fig. 5 maps the spatial distribution of the steady-state per-capita income levels predicted by our estimated WAC model. It can be seen that the steady-state incomes are markedly higher in the main regional urban centres, Hasselt and Sint-Truiden, along with their less densely populated exurban and rural-urban fringes, than in the rest of the province. By portraying the express highways and (main) railway stations in the province, the map also reveals that all municipalities in the high/middle-high steady-state income group are ‘infrastructure-rich’. Given their location at close proximity to express highways and railway stations (providing rush-hour service to other cities in Belgium), it is also conceivable that many high-income residents of these municipalities out-commute over longer distances to their workplaces outside the province of Limburg (e.g., Brussels or Antwerp), where wages are generally known to be higher.

[Fig. 5 about here]

## SUMMARY AND CONCLUSIONS

This paper represents an attempt to shed some light on the complex web of economic-geographic pattern of migration in the Belgian province of Limburg and its effect in terms of the (spatial) re-distribution of taxable incomes per capita across municipalities.

Using cross-sectional data at the municipality level, it has been shown that the 1990s was a period of substantial  $\sigma$ -divergence. We also found that growth rates of per-capita incomes were not randomly distributed over space. Based on a model, built on the notion of weak absolute convergence (WAC) and incorporating both observed and unobserved heterogeneity, we found that per-capita incomes exhibited varying rates of growth and  $\beta$ -convergence, with an average convergence rate of about 4.2% per annum. Furthermore, and more importantly, we found that both migration and education turned out to be nurturing cross-sectional income dispersion or  $\sigma$ -divergence, with the highest-income municipalities being spatially clustered in the south-western part of the province – i.e., both in and around the major regional urban centres of Hasselt (capital) and Sint-Truiden, at the expense of the municipalities that are

mainly located in the north-eastern part of the province. This, in turn, led us to conclude that the spatial distribution of service-based amenities and other quality-of-life factors do matter in directing the residence and migration patterns of high-income people.

The results suggest that the municipalities that entered the 1990s with the highest per-capita incomes were those that performed comparatively well due to the influx of high-income migrants. The finding that the migratory process favoured the initially 'rich' municipalities more than the other municipalities provides indirect support for the hypothesis of selective income migration. Due to the presence of spatial spill-over effects, also the immediately neighbouring municipalities tended to have high income growth rates, as they were also able to attract the highest-income migrants and as local non-migrants were able to considerably improve their income positions due to the strengthened economic base of these municipalities.

Based on our empirical results, it appears as if high-income people sort themselves into specific locations based on their tastes for amenities and other quality-of-life factors. In other words, there seems to be a tendency of high-income people selecting to reside near people with similar income positions. Given that amenities are, at least to some extent, endogenous, this may provide the basis for a continuing process of cumulative causation, since the pull effect of local amenities is likely to strengthen in the future.

The divergence of per-capita incomes and the spatial concentration of high-income municipalities in the province may have important policy implications, given its potential impact on local tax bases, local housing prices, local demand for both private and public services, firm locations (e.g., particular types of firms/jobs may 'follow' the well-educated, high-income people), and so on. However, since migration is a multi-dimensional issue, with multiple causes and effects, it is difficult to make firm policy recommendations based on the present research. Yet, the foregoing study inevitably brings us to the following question: If it is believed that widening income disparities between municipalities are undesirable, can this process be reverted by public policies? In answering this question, it is important that policies designed to reduce or revert income disparities should at least take into account the incidence of selective income migration to amenity-rich areas. One possible option policy-makers could take is, for example, 'to unlock' the peripheral areas in the north-eastern part of the province by targeting more investments in public transportation infrastructures (for example, the Light-

Rail network and the North-South express highway), and 'to tear down the border' by promoting interactions with major urban centres in the Dutch province of Limburg (for example, Eindhoven in the north, and Sittard-Geleen/Maastricht in the east).

## APPENDIX: GME FORMULATION OF THE EMPIRICAL MODEL

To implement the GME method (GOLAN *et al.*, 1996; FRASER, 2000), the estimation problem of the WAC model has to be converted into a constrained optimization problem, where the objective function consists of the joint entropy in equation (A.1) below. This objective function is to be maximized, subject to the appropriate data-consistency and normalization constraints.

### *Re-parameterization*

The parameter set  $\eta = \{\alpha, \gamma_0, \gamma_1, \gamma_2\}$  of the empirical WAC model in equation (6), along with the spatial autocorrelation parameter  $\lambda$  and the error term  $u_i$  in equation (2), can be re-parameterized in terms of a set of unknown probability vectors  $\mathbf{p}_\eta = (p_{\eta,1}, \dots, p_{\eta,M})'$ ,  $\mathbf{p}_\lambda = (p_{\lambda,1}, \dots, p_{\lambda,M})'$ , and  $\mathbf{p}_{u_i} = (p_{u_i,1}, \dots, p_{u_i,G})'$ , respectively, and the corresponding support vectors  $\mathbf{s}_\eta = (s_{\eta,1}, \dots, s_{\eta,M})'$ ,  $\mathbf{s}_\lambda = (s_{\lambda,1}, \dots, s_{\lambda,M})'$ , and  $\mathbf{s}_u = (s_{u,1}, \dots, s_{u,G})'$ . The parameters and the error term are then re-parameterized as  $\eta = \mathbf{p}'_\eta \mathbf{s}_\eta$ ,  $\lambda = \mathbf{p}'_\lambda \mathbf{s}_\lambda$ , and  $u_i = \mathbf{p}'_{u_i} \mathbf{s}_{u_i}$ .

### *Optimization problem*

The GME formulation of the empirical WAC model is

$$\text{Max}_{\mathbf{p}} H(\mathbf{p}) = -\mathbf{p}'_\alpha \ln(\mathbf{p}_\alpha) - \mathbf{p}'_{\gamma_0} \ln(\mathbf{p}_{\gamma_0}) - \mathbf{p}'_{\gamma_1} \ln(\mathbf{p}_{\gamma_1}) - \mathbf{p}'_{\gamma_2} \ln(\mathbf{p}_{\gamma_2}) - \mathbf{p}'_\lambda \ln(\mathbf{p}_\lambda) - \sum_{i=1}^N \mathbf{p}'_{u_i} \ln(\mathbf{p}_{u_i}) \quad (\text{A.1})$$

subject to

$$g_i - \left\{ \mathbf{p}'_\lambda \mathbf{s}_\lambda \sum_{j=1}^N \tilde{w}_{ij} \left( \frac{g_j - \mathbf{p}'_\alpha \mathbf{s}_\alpha - (\mathbf{p}'_{\gamma_0} \mathbf{s}_{\gamma_0} + \mathbf{p}'_{\gamma_1} \mathbf{s}_{\gamma_1} m_j + \mathbf{p}'_{\gamma_2} \mathbf{s}_{\gamma_2} x_{j0}^{-1}) y_{j0}}{y_{j0}} \right) \right\} y_{i0} \quad (\text{A.2})$$

$$= \mathbf{p}'_\alpha \mathbf{s}_\alpha + (\mathbf{p}'_{\gamma_0} \mathbf{s}_{\gamma_0} + \mathbf{p}'_{\gamma_1} \mathbf{s}_{\gamma_1} m_i + \mathbf{p}'_{\gamma_2} \mathbf{s}_{\gamma_2} x_{i0}^{-1}) y_{i0} + \mathbf{p}'_{u_i} \mathbf{s}_u y_{i0}, \quad i = 1, \dots, N$$

where

$$\delta_i = \mathbf{p}'_\lambda \mathbf{s}_\lambda \sum_{j=1}^N \tilde{w}_{ij} \left( \frac{g_j - \mathbf{p}'_\alpha \mathbf{s}_\alpha - (\mathbf{p}'_{\gamma_0} \mathbf{s}_{\gamma_0} + \mathbf{p}'_{\gamma_1} \mathbf{s}_{\gamma_1} m_j + \mathbf{p}'_{\gamma_2} \mathbf{s}_{\gamma_2} x_{j0}^{-1}) y_{j0}}{y_{j0}} \right) + \mathbf{p}'_{u_i} \mathbf{s}_{u_i}, \quad i = 1, \dots, N \quad (\text{A.3})$$

and

$$\sum_{i=1}^N \delta_i = 0 \quad (\text{A.4})$$

$$\begin{aligned} \sum_{m=1}^M p_{\alpha,m} = 1; \quad \sum_{m=1}^M p_{\gamma_0,m} = 1; \quad \sum_{m=1}^M p_{\gamma_1,m} = 1; \quad \sum_{m=1}^M p_{\gamma_2,m} = 1; \quad \sum_{m=1}^M p_{\lambda,m} = 1; \\ \sum_{g=1}^G p_{u_i,g} = 1, \quad i = 1, \dots, N \end{aligned} \quad (\text{A.5})$$

Equation (A.1) denotes the entropy objective, which is subject to the data-consistency constraints in equation (A.2). Equation (A.3) defines the unobserved effects, while the constraints in equation (A.4) preserve a mean unobserved effect of zero. Finally, the constraints in equation (A.5) ensure that all unknown probabilities add up to one.

After solving the entropy optimization problem in (A.1) through (A.5), the parameter estimates and the error terms can be recovered as  $\hat{\boldsymbol{\eta}} = \hat{\mathbf{p}}'_\eta \mathbf{s}_\eta$ ,  $\hat{\lambda} = \hat{\mathbf{p}}'_\lambda \mathbf{s}_\lambda$ , and  $\hat{u}_i = \hat{\mathbf{p}}'_{u_i} \mathbf{s}_{u_i}$ , respectively, where  $\hat{\mathbf{p}}_\eta$ ,  $\hat{\mathbf{p}}_\lambda$ , and  $\hat{\mathbf{p}}_{u_i}$  are the corresponding estimated probabilities. The implied  $\hat{\delta}_i$ 's can then easily be derived by using equation (5) in the main text.

### *Support ranges*

For estimation purposes, a common support vector for the coefficients in  $\boldsymbol{\eta}$  is used. Since we have little prior knowledge about the ‘true’ value of the coefficient, the support vector is set as  $\mathbf{s}_\eta = (-100, -50, 0, 50, 100)'$ , which represents a range wide enough to include all possible outcomes (see also, for example, GOLAN *et al.*, 2001). On the other hand, the spatial autocorrelation coefficient  $\lambda$  should normally lie within the  $[-1, 1]$  interval. Hence, its support vector is defined as  $\mathbf{s}_\lambda = (-1, -0.5, 0, 0.5, 1)'$ . Finally, the error supports are defined as in GOLAN *et al.* (1996), based on PUKELSHEIM’s (1994) ‘three-sigma’ rule. Hence, the support vector for each individual  $u_i$  is set as  $\mathbf{s}_{u_i} = (-3\sigma_g/y_{i0}, 0, 3\sigma_g/y_{i0})'$ , where  $\sigma_g$  is the



empirical standard deviation of the per-capita income-growth rates (i.e., the dependent variable in the growth regression).

## REFERENCES

- ADAMSON D.W., CLARK D.E. and PARTRIDGE M.D. (2004) Do urban agglomeration effects and household amenities have a skill bias? *Journal of Regional Science* 44, 201-223.
- ANJOMANI A. (2002) Regional growth and interstate migration. *Socio-Economic Planning Sciences* 36, 239-265.
- ALONSO W. (1964) *Location and Land Use*. Harvard University Press, Cambridge, MA.
- ARONSSON T., LUNDBERG J. and WIKSTRÖM M. (2001) Regional income growth and net migration in Sweden, 1970-1995. *Regional Studies* 35, 823-830.
- BARRO R. J. and SALA-I-MARTIN X. (2003) *Economic Growth: Second Edition*. The MIT Press.
- BJØRN E., LINDQUIST K.-G. and SKJERPEN T. (2003) Random coefficients in unbalanced panels: an application on data from chemical plants. *Annales d'Économie et de Statistique* 69, 55-83.
- BOYLE G.E. and MCCARTHY T.G. (1997) A simple measure of  $\beta$ -convergence. *Oxford Bulletin of Economics and Statistics* 59, 257-264.
- BRUECKNER J.K., THISSE J.-F. and ZENOU Y. (1999) Why is central Paris rich and downtown Detroit poor? An amenity-based theory. *European Economic Review* 43, 91-107.
- BURNLEY I.H., MURPHY P.A. and JENNER A. (1997) Selecting suburbia: residential relocation to outer Sydney. *Urban Studies* 34, 1109-1127.
- CANOVA F. (1999) *Testing for Convergence Clubs in Income Per Capita: A Predictive Density Approach*. CEPR Working Paper No. 2201, London.
- CANOVA F. and MARCET A. (1995) *The Poor Stay Poor: Non-Convergence Across Countries and Regions*. CEPR Discussion Paper No. 1265, London.
- CARRIÓN-FLORES C. and IRWIN E.G. (2004) Determinants of residential land-use conversion and sprawl at the rural-urban fringe. *American Journal of Agricultural Economics* 86, 889-904.
- CROMARTIE J.B. and NORD M. (1997) Migration contributes to nonmetro per capita income growth. *Rural Conditions and Trends* 8, 40-45.
- CUSHING B. and POOT J. (2004) Crossing boundaries and borders: regional science advances in migration modeling. *Papers in Regional Science* 83, 317-338.

- DELLER S.C., TSAI T.H., MARCOUILLER W. and ENGLISH D.B.K. (2001) The role of amenities and quality of life in rural economic growth. *American Journal of Agricultural Economics* 83, 352-365.
- DESDOIGTS A. (1999) Patterns of economic development and the formation of clubs. *Journal of Economic Growth* 4, 305-330.
- DEVOGELAER D. (2004) *Interne migraties in België: wie, waarom en naar welke gemeenten? (Internal migrations in Belgium: who, why and to which municipalities?)* Working Paper No. 5-04, Federal Planning Bureau, Brussels.
- DEWHURST J.H.L. and MUTIS-GAITAN H. (1995) Varying speeds of regional GDP per capita convergence in the European Union, 1981-91. In H.W. Armstrong and R.W. Vickerman (eds.), *Convergence and Divergence Among European Regions*. European Research in Regional Science, 5, pp. 22-39.
- DRENNAN M.P. and LOBO J. (1999) A simple test for convergence of metropolitan income in the United States. *Journal of Urban Economics* 46, 350-359.
- DURLAUF S.N. and JOHNSON P.A. (1995) Multiple regimes and cross-country growth behaviour. *Journal of Applied Econometrics* 10, 365-384.
- EFRON B. and TIBSHIRANI R. (1993) *An Introduction to the Bootstrap*. Monographs on Statistics and Applied Probability, 57. Chapman & Hall, New York.
- ELHORST J.P. (2003) Specification and estimation of spatial panel data models. *International Regional Science Review* 26, 244-268.
- EVANS P. (1997) How fast do economies converge? *The Review of Economics and Statistics* 79, 219-225.
- FOTHERINGHAM A.S. and O'KELLY M.E. (1989) *Spatial Interaction Models: Formulations and Applications*. Kluwer Academic Publishers, Dordrecht.
- FRASER I. (2000) An application of maximum entropy estimation: the demand for meat in the United Kingdom. *Applied Economics* 32, 45-59.
- FRIEDBERG R.M. and HUNT J. (1995) The impact of immigrants on host country wages, employment and growth. *Journal of Economic Perspectives* 9(2), 23-44.
- GOLAN A., JUDGE G. and MILLER D. (1996) *Maximum Entropy Econometrics: Robust Estimation with Limited Data*. New York: John Wiley & Sons.

- GOLAN A., PERLOFF J.M. and SHEN Z. (2001) Estimating a demand system with non-negativity constraints: Mexican meat demand. *The Review of Economics and Statistics* 8, 541-550.
- GOTTLIEB P.D. (1993) Residential amenities, firm location and economic development. *Urban Studies* 32, 1413-1436.
- GREENWOOD M.J. (1985) Human migration: theory, models, and empirical studies. *Journal of Regional Science* 25, 521-544.
- GREENWOOD M.J. and HUNT G.L. (1989) Jobs versus amenities in the analysis of metropolitan migration. *Journal of Urban Economics* 25, 1-16.
- HIGGINS M.J., LEVY D. and YOUNG A.T. (2006) Growth and convergence across the United States: evidence from county-level data. *The Review of Economics and Statistics* 88, 671-681.
- IRWIN E.G. and BOCKSTAEEL N.E. (2002) Interacting agents, spatial externalities and the evolution of residential land use patterns. *Journal of Economic Geography* 2(2), 31-54.
- ISLAM N. (2003) What have we learnt from the convergence debate? *Journal of Economic Surveys* 17, 309-362.
- KANGASHARJU A. (1998)  $\beta$  convergence in Finland: regional differences in speed of convergence. *Applied Economics* 30, 679-687.
- KNAPP T.A. and GRAVES P.H.E. (1989) On the role of amenities in models of migration and regional development. *Journal of Regional Science* 29, 71-87.
- LAURINI M., ANDRADE E. and VALLS PEREIRA P.L. (2005) Income convergence clubs for Brazilian municipalities: a non-parametric analysis. *Applied Economics* 37, 2099-2118.
- LENCE S.H. and MILLER D.J. (1998) Recovering output-specific inputs from aggregate input data: a generalized cross-entropy approach. *American Journal of Agricultural Economics* 80, 852-867.
- MADDALA G.S and WU S. (2000) Cross-country growth regressions: problems of heterogeneity, stability and interpretation. *Applied Economics* 32, 635-642.
- MARGO R.A. (1992) Explaining the postwar suburbanization in the United States: the role of income. *Journal of Urban Economics* 31, 300-310.
- MATHUR V.K. and STEIN S.H. (2005) Do amenities matter in attracting knowledge workers for regional economic development? *Papers in Regional Science* 84, 251-269.

- MILLS E.S. (1967) An aggregative model of resource allocation in a metropolitan area. *American Economic Review* 57, 197-210.
- MUESER P.R. and GRAVES P.E. (1995) Examining the role of economic opportunity and amenities in explaining population redistribution. *Journal of Urban Economics* 37, 176-200.
- MUTH R.F. (1969) *Cities and Housing*. University of Chicago Press, Chicago.
- PLANE D.A. (1999) Geographical pattern analysis of income migration in the United States. *International Journal of Population Geography* 5, 195-212.
- PLANE D.A. and BITTER C. (1997) The role of migration research in regional science. *Papers in Regional Science* 76, 133-153.
- PLANE D.A. and HEINS F. (2003) Age articulation of U.S. inter-metropolitan migration flows. *Annals of Regional Science* 37, 107-130.
- PORELL F.W. (1982) Intermetropolitan migration and quality of life. *Journal of Regional Science* 22, 137-158.
- PUKELSHEIM F. (1994) The three-sigma rule. *The American Statistician* 48, 88-91.
- RENKOW M. and HOOVER D. (2000) Commuting, migration, and rural-urban population dynamics. *Journal of Regional Science* 40, 261-287.
- ROBERTS M. (2004) The growth performances of the GB counties: some new empirical evidence for 1977-1993. *Regional Studies* 38, 149-165.
- SÁ C., FLORAX R.J.G.M. and RIETVELD P. (2004) Determinants of the regional demand for higher education in The Netherlands: a gravity model approach. *Regional Studies* 38, 375-392.
- SHELLEY M.C. and KOVEN S.G. (1993) Interstate migration: a test of competing interpretations. *Policy Studies Journal* 21, 243-261.
- SHIOJI E. (2001) Composition effect of migration and regional growth in Japan. *Journal of the Japanese and International Economies* 15, 29-49.
- SHUMWAY J.M. and OTTERSTROM S.M. (2001) Spatial patterns of migration and income change in the Mountain West: the dominance of service-based, amenity-rich counties. *Professional Geographer* 53, 492-502.
- SIMPSON W. and VAN DER VEEN A. (1992) The economics of commuting and the urban labour market. *Journal of Economic Surveys* 6, 45-62.

VAN HECKE E. (1998) Actualisering van de stedelijke hiërarchie in België (Updating the urban hierarchy in Belgium). *Het Tijdschrift van het Gemeentekrediet* 52(205), 45-76.

Table 1. Descriptive statistics for per-capita taxable incomes

	1991	1993	1995	1997	1999	2000
Mean	7.915	8.178	8.858	9.177	9.695	10.122
Median	7.914	8.103	8.871	9.196	9.741	10.048
Maximum	9.583	9.762	10.587	10.950	11.525	12.026
Minimum	6.977	7.137	7.591	7.908	8.096	8.090
Standard deviation	0.515	0.525	0.612	0.655	0.684	0.961
Coefficient of variation (Normalized C.V.)	0.065 (1.000)	0.064 (0.988)	0.069 (1.063)	0.071 (1.098)	0.071 (1.086)	0.095 (1.460)
Skewness	0.649	0.655	0.313	0.377	0.076	-0.043
Kurtosis	1.246	0.832	0.529	0.388	0.593	-0.423
Kendall's <i>W</i>	1.000	0.976	0.945	0.927	0.921	0.904
Spearman's <i>rho</i>	1.000	0.948	0.887	0.906	0.898	0.760

Note: Per-capita taxable incomes are measured in 1,000 EUROS, and deflated/inflated using the consumer's price index (constant 1996 prices).

Table 2. Descriptive statistics for net in-migration and education (percentages)

	Average annual net in-migration rate, 1991-2000	Initial level of higher education, 1991
Mean	0.28%	21.1%
Standard deviation (Coefficient of variation)	0.28% (102.7%)	3.1% (14.6%)
Median	0.33%	21.2%
Minimum	-0.62%	15.2%
Maximum	0.92%	31.0%

Note: The annual net in-migration rate in each municipality is measured per 100 inhabitants. The initial level of higher education is defined as the percentage of the local working population with a post-secondary or university degree.



Table 3. Estimation results from the GME estimation of the WAC model

	Panel A: Point estimates, spatial weighting based on <i>contiguity</i> (Bootstrap standard errors) [Bootstrap 90% confidence intervals]	Panel B: Point estimates, spatial weighting based on <i>distance</i> (Bootstrap standard errors) [Bootstrap 90% confidence intervals]
$\hat{\alpha}$ (Constant, general)	0.1226 (0.0313) [0.0693; 0.1752]	0.0995 (0.0321) [0.0460; 0.1450]
$\hat{\gamma}_0$ (Constant, in varying $\beta_i$ )	-0.0361 (0.0135) [-0.0646; -0.0137]	-0.0260 (0.0136) [-0.0448; -0.0037]
$\hat{\gamma}_1$ (Net in-migration rate)	0.2721 (0.1384) [0.0699; 0.5068]	0.2537 (0.1702) [0.0166; 0.5839]
$\hat{\gamma}_2$ (Education inverse)	-0.0023 (0.0008) [-0.0037; -0.0010]	-0.0020 (0.0008) [-0.0035; -0.0006]
$\hat{\lambda}$ (Spatial correlation)	0.6652 (0.0633) [0.5428; 0.7303]	0.6559 (0.0451) [0.5699; 0.7086]
Unobserved effects ( $\hat{\delta}_i$ )		
Mean	0.0000	0.0003
Standard deviation	0.0015	0.0017
Minimum	-0.0029	-0.0027
Maximum	0.0031	0.0031
$\beta$ -coefficients ( $\hat{\beta}_i$ )		
Mean	-0.0463	-0.0348
Standard deviation	0.0025	0.0025
Minimum	-0.0529	-0.0407
Maximum	-0.0420	-0.0309
Annual convergence rates ( $\hat{r}_i$ )		
Mean	6.00	4.18
Standard deviation	0.43	0.36
(Coefficient of variation)	(7.2%)	(8.6%)
Minimum	5.28	3.62
Maximum	7.19	5.08
Pseudo- $R^2$	0.47	0.54

Note: The results are based on Generalized Maximum Entropy estimation (see Appendix), using the GAMS software package (CONOPT3 solver).

Fig. 1. Moran scatter plots of per-capita incomes, 1991 and 2000

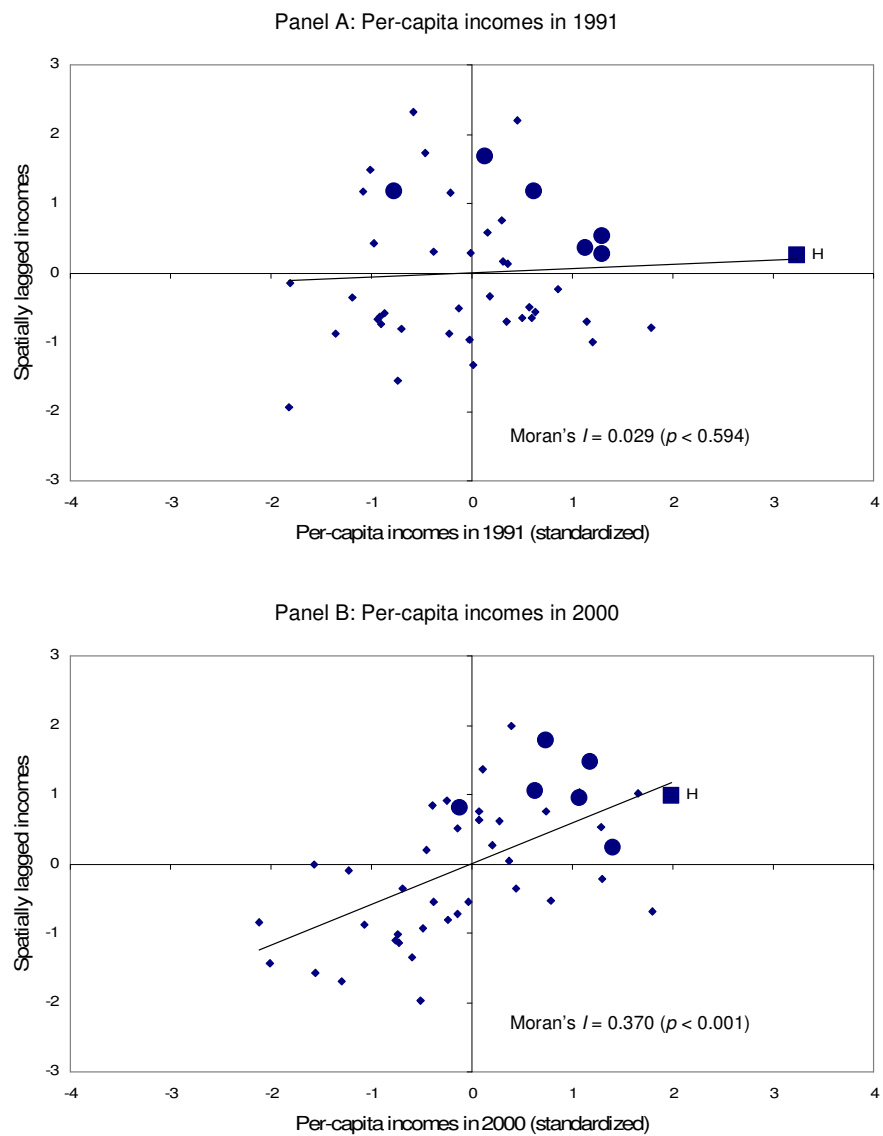
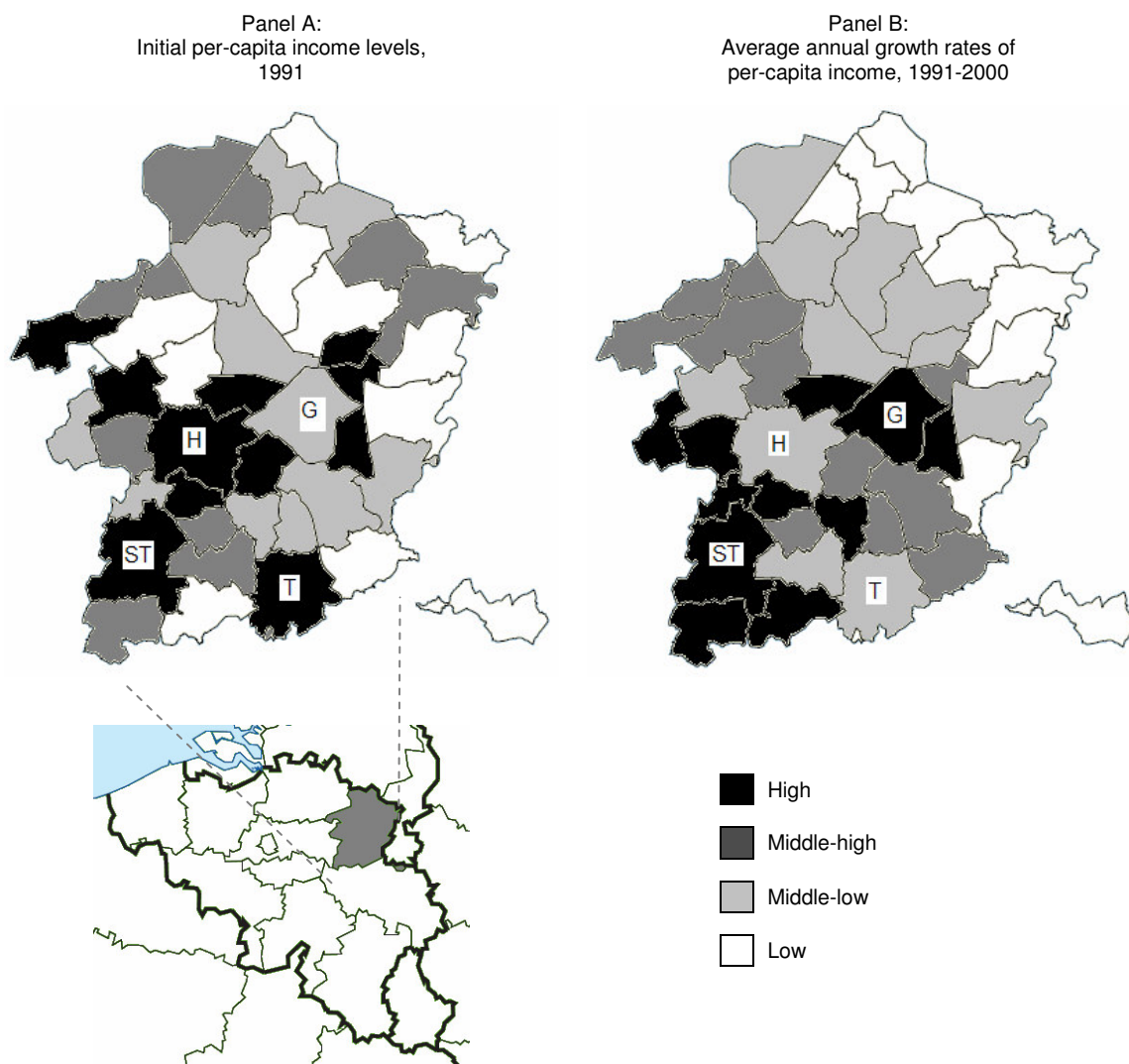


Fig. 2. Spatial distribution of initial per-capita income levels, 1991, and average annual growth rates of per-capita income, 1991-2000 (quartiles)



Note: H = Hasselt (capital), G = Genk, ST = Sint-Truiden, and T = Tongeren are the main regional urban centres in the province of Limburg, all having a strong amenity advantage (VAN HECKE, 1998).

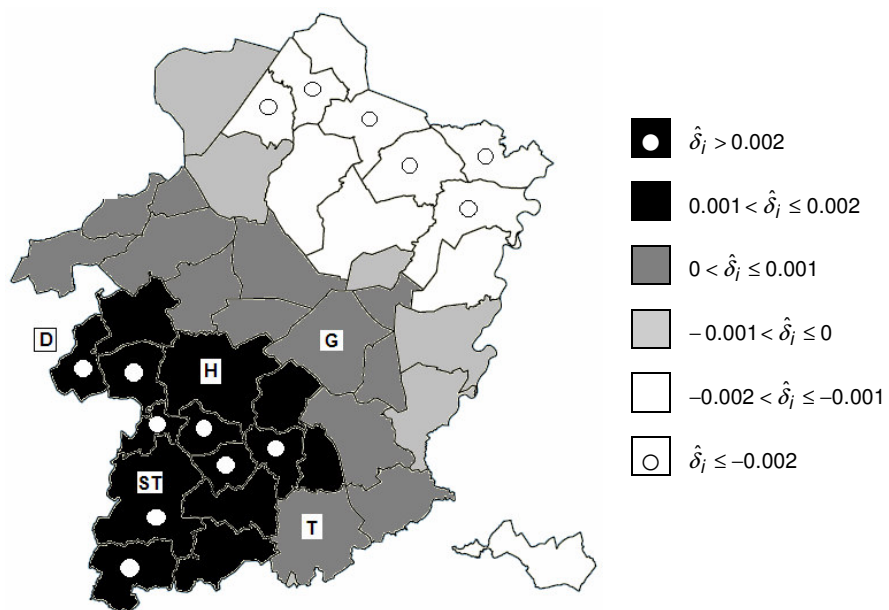
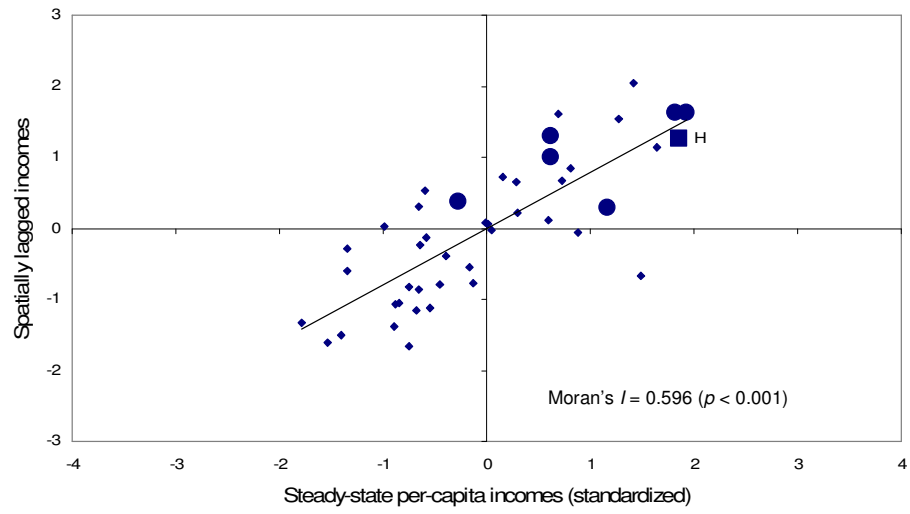
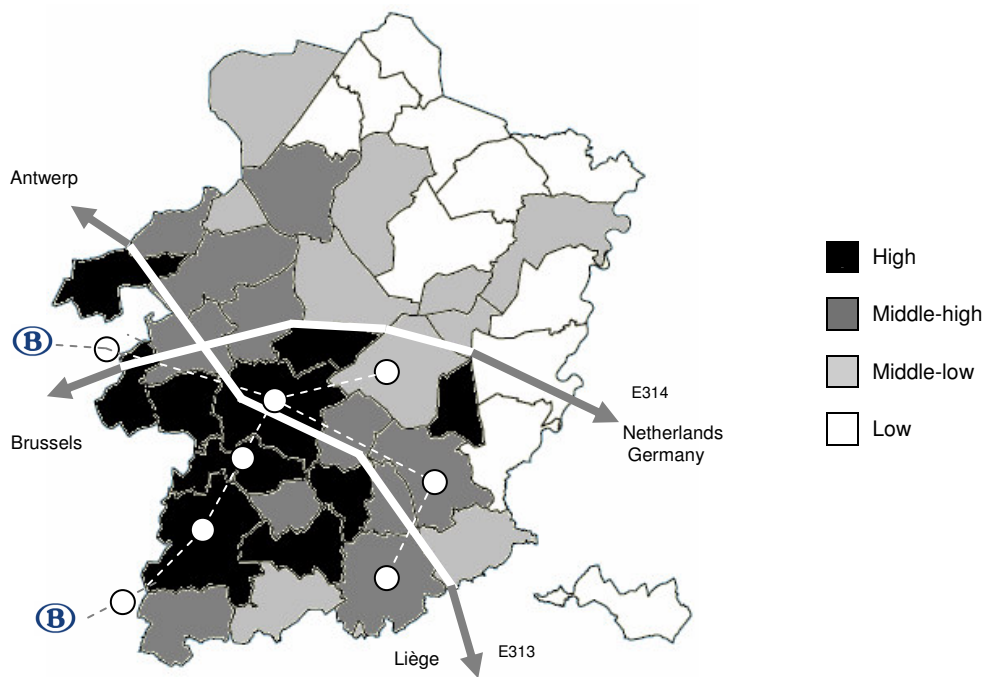
Fig. 3. Spatial distribution of estimated unobserved effects ( $\delta_i$ )

Fig. 4. Moran scatter plot of predicted steady-state per-capita incomes



Note: Hasselt (H) and its neighbouring municipalities are marked as ■ and ●, respectively.

Fig. 5. Spatial distribution of predicted steady-state per-capita incomes (quartiles)



## NOTES

---

<sup>1</sup> By ‘amenities’ we mean a myriad of public and private services, such as administrative agencies, schools, hospitals, shops, restaurants, entertainment and recreation facilities, etc. Amenities may also include access to express highways and (rush-hour) train service.

<sup>2</sup> In the present paper, we are dealing with *total* migration, not just labour mobility, such as in GREENWOOD and HUNT (1989), and SIMPSON and VAN DER VEEN (1992), among others. Using total migration data may be one of the reasons for the mixed results found in the literature concerning the determinants of migration.

<sup>3</sup> It has to be recognized, though, that if jobs are attractive to migrants, and if jobs are growing most rapidly in amenity-rich areas, then to some extent, at least, the presence of amenities may be reflecting the importance of job opportunities (GREENWOOD, 1985). Exploratory research, however, revealed that (paid) employment was *not* growing most rapidly in amenity-rich locations (see also footnote 8). Moreover, one may reasonably suspect that, in the present context of the province of Limburg, residential migration dominates labour migration, given the (relative) immobility of the labour force in Belgium, the small size of the region, and the presence of an excellent transportation network (i.e., low commuting costs). Hence, many people may change their place of residence, without changing their workplace.

<sup>4</sup> Limburg is the easternmost province of Flanders, bordering on the Netherlands (see Fig. 2). It is a small province, with a total population of about 800,000, covering an area of 2,414 square kilometres. Limburg is an interesting case, as it forms some kind of ‘microcosm’: within the context of Flanders, it is, relatively speaking, a ‘remote’ region, with no large metropolitan areas, but with a few ‘regional urban centres’. Hasselt is the largest centre, with about 70,000 inhabitants; it is the capital of the province and, along its twin (industrial) city of Genk, the main traffic artery in the region. The rest of the province is largely characterized by (easily accessible) small residential towns and rural-type areas.

<sup>5</sup> Taxable income is defined as the pre-tax income received by households, net of social-security contributions; it includes wages, salaries and other labour incomes, dividends,

---

interests and rents, proprietor's net incomes, and transfer payments. Per-capita taxable income is defined as the ratio of total taxable income in a municipality to the resident population.

<sup>6</sup> It can safely be assumed that price-level disparities between municipalities in a small region such as the province of Limburg are non-existent. As a result, applying the same price deflator to all municipalities means that, in effect, we are still only comparing nominal figures.

<sup>7</sup> The industrial city of Genk (a former mine town) takes a somewhat different position, due to the relatively high proportion of low-income immigrants (ethnic minorities) in the total population.

<sup>8</sup> It should be noted that we were not able to find any positive relationship between income-growth rates and employment growth rates in the province; the correlation coefficients between the income-growth rates and the growth rates of total employment and employment in knowledge-intensive industries are -0.336 ( $p = 0.026$ ) and -0.157 ( $p = 0.307$ ), respectively. We interpret these findings as supporting the view that the pattern of income changes is primarily related to the spatial distribution of amenities (see also footnote 3).

<sup>9</sup> In using the term 'weak absolute convergence', we took inspiration from ISLAM (2003, p. 330), who introduced the notion of weak *conditional* convergence to indicate a process (within the framework of the neo-classical growth theory) whereby economies are converging to *different* levels – or growth paths – of per-capita income, but at different rates. Here, we assume that per-capita incomes converge to a *common* balanced-growth path, but at different speeds.

<sup>10</sup> Despite the fact that the issue of heterogeneity has not gone unnoticed in the convergence literature, there are only a few empirical studies explicitly dealing with it. For example, DEWHURST and MUTIS-GAITAN (1995) and KANGASHARJU (1998) adopted a cross-sectional approach, and allowed for the possibility that regions converge to equilibrium growth rates at different speeds (by using a varying-coefficients model that is, to some limited extent, similar



---

to the one developed in this paper). The idea of heterogeneity in convergence rates is also linked to the vast literature dealing with ‘club convergence’ or ‘multiple regimes’ (e.g., DURLAUF and JOHNSON, 1995; CANOVA, 1999; DESDOIGTS, 1999; LAURINI *et al.*, 2005). Although this literature acknowledges the potential differences in the speeds of convergence, any (residual) heterogeneity within the identified sub-groups is still ruled out. Consequently, the regression coefficients for each sub-group still represent averages of the underlying individual coefficients for each observational unit. On the other hand, CANOVA and MARCET (1995) and MADDALA and WU (2000) used (empirical or hierarchical) Bayesian methods to account for heterogeneity in convergence rates, but they do so by using a panel-data, rather than a cross-sectional, framework. Importantly, both papers demonstrated substantial biases in convergence rates under the assumption of parameter constancy or homogeneity.

<sup>11</sup> Using a model with parameter constancy (homogeneity) would only provide a consistent estimate of a single  $\beta$  if the varying coefficients differ *randomly* and be distributed *independently* of initial income  $y_{i0}$ .

<sup>12</sup> The intra-municipal distance is defined as  $d_{ii} = ((\pi - 1)/\pi)\sqrt{s_i/\pi}$ , where  $s_i$  is the surface area (in square kilometres) of municipality  $i$  (see, for example, SÁ *et al.*, 2004).

<sup>13</sup> The notion of centrality index has been borrowed from the literature on spatial-interaction or gravity models (see e.g., FOTHERINGHAM and O’KELLY, 1989; SÁ *et al.*, 2004).

<sup>14</sup> Some authors (e.g., HIGGINS *et al.*, 2006) have attempted to include a host of explanatory variables in the growth regression to account for cross-sectional heterogeneity. However, in the present context (using cross-sectional data at the municipal level), such an approach is considered inappropriate (given the many interactions between the observational units) and infeasible (due to data limitations and/or measurement problems).

<sup>15</sup> It should be noted that the data on net in-migration include both ‘internal’ and ‘external’ migration. In 2000, for example, about 35% (on average) of the net in-flux of migrants

---

originated from outside of the province (either from other Belgian regions or from abroad). As a result, the sum (or the average) of the net in-migration rates over the municipalities is not equal to zero, and almost all municipalities in the province experienced a positive net in-migration in the 1990s.

<sup>16</sup> This advantage holds because we use cross-sectional data. If we would have used panel data, the model could be estimated along the lines spelled out in, for example, ELHORST (2003). In a panel-data framework, however, the individual estimates  $\delta_i$  (one of our primary concerns in this paper) would have been differenced out.

<sup>17</sup> It should be noted that the estimation results may be sensitive to the prior information. Specifically, the final GME estimates may be ‘pulled’ towards the centre of the support range (e.g., LENCE and MILLER, 1998). Therefore, we tested for this sensitivity by applying a ‘quasi-Bayesian’ procedure. This procedure involves the use of randomized prior values of the parameters, drawn from a ‘vague’ uniform prior distribution, and the minimization of the cross-entropy criterion (for details on the cross-entropy criterion, see GOLAN *et al.*, 1996, and FRASER, 2000). The results of this procedure, which are available from the author upon request, indicated that the outcomes are fairly robust – that is, the posterior means are very close to the final estimates reported in this paper.

<sup>18</sup> Since the (small) sampling properties of the GME estimator are generally unknown, we employed a simple percentile bootstrap approach (EFRON and TIBSHIRANI, 1993) to assess the precision of the GME estimator. The bootstraps are derived by using a re-sampling procedure rather than an ‘error bootstrap’, to preserve the spatial structure of the data. We randomly took 200 samples from the empirical joint distribution of the data (i.e., from the original sample), and formed 200 new data matrices. Then, we re-estimated the model 200 times using the newly-formed data matrices, yielding an equal number of different bootstrap estimates. The observed distributions of these estimates were then taken as approximations of the ‘true’ distribution of the estimates, from which 90% confidence intervals are calculated.

---

<sup>19</sup> Even though the magnitudes of some of the estimates seem to be sensitive to the type of spatial weighting, the conclusions presented in this paper broadly remain the same.

<sup>20</sup> It may be argued that the net in-migration variable  $m_i$  is potentially endogenous, and that an instrumental-variable (IV) approach would be more appropriate to estimate the (quantity) effect of net in-migration, due to the fact that some omitted variables may possibly affect both per-capita income growth and net in-migration in a similar way (i.e., factors not controlled for in the model may give rise to higher per-capita income growth *and* higher net in-migration). Though this could perhaps be the case, some experimentation involving IV techniques (where we instrumented net in-migration by using the predicted values from an OLS regression of a migration equation) demonstrated that the estimates are quite robust both in sign and magnitude. Such a finding is not surprising, given the fact that we already control for unobserved factors by including the location-specific terms  $\delta_i$  in the WAC model. Moreover, given the small number of observations, using IV may do more harm than good (e.g., small-sample bias may sneak in) and usually goes with a cost in terms of higher standard errors and wider confidence intervals (which may hamper proper statistical inference). Therefore, we decided to report only the ‘non-IV’ results in Table 3.

<sup>21</sup> The significance of  $\hat{\lambda}$  indicates that not controlling for unobserved effects would lead to omitted-variable bias, whilst the correlation between the  $\hat{\delta}_i$ s and the initial-income variable (which is equal to 0.296,  $p = 0.051$ ) indicates that these unobserved effects should be modelled as fixed effects rather than random effects (e.g., BJØRN *et al.*, 2003). Also, using the centrality indexes to account for the unobserved cross-sectional heterogeneity removes any remaining (global) spatial autocorrelation. For example, the Moran’s  $I$  for the residuals  $\hat{u}_i$  is equal to 0.008 ( $p = 0.750$ ).

<sup>22</sup> Evidently, unobserved effects may spill over from beyond the (internal) border of the province. This may particularly be the case for the westernmost municipalities, which are located at close proximity to the regional urban centre of Diest (D) (see Fig. 3).

---

<sup>23</sup> In a way, this phenomenon somehow reconciles two opposing perspectives in the urban-economics literature: the amenity-based theory (e.g., BRUECKNER *et al.*, 1999) *versus* the mono-centric city model (e.g., ALONSO, 1964; MILLS, 1967; MUTH, 1969).

<sup>24</sup> Unfortunately, it is not possible to assess the change over the study period in the educational attainment of the working population (data were only available for 1991).

<sup>25</sup> Note further that the speed of convergence to this steady-state income is  $(1 - \rho_i^{1/T}) = [1 - (1 + \beta_i)^{1/T}]$  per year, and that the convergence rates approach zero as  $T$  goes to infinity.