

Paper to be presented at the ERSA 41st European Congress in Zagreb, Croatia
29th August - 1st September 2001

CAN MIGRATION DECISIONS BE AFFECTED BY INCOME TAXATION POLICIES?

Mika Haapanen*
University of Jyväskylä
School of Business and Economics

[Email: mphaapan@tase.jyu.fi](mailto:mphaapan@tase.jyu.fi)

Jari Ritsilä
University of Jyväskylä
School of Business and Economics
Centre for Economic Research

[Email: jarrits@tase.jyu.fi](mailto:jarrits@tase.jyu.fi)

4th June 2001

Mailing Address: University of Jyväskylä, School of Business and Economics,
PO Box 35, FIN-40351 Jyväskylä, Finland.

Abstract

The question of individually focused policy measures is currently at the centre of public debate in Finland. Is it possible through income taxation policy to influence where people choose to live? This paper analyses the role which expected net income might play in the redistribution of population. Random parameter logit model is used in the empirical analyses of migration. Micro-level analyses of the Finnish data show that the decision to migrate is influenced by expected net income, but the size of the impact depends on age, home-ownership and the regional unemployment rate. The estimation results also imply that a reduction in income taxation would be most effective if it were targeted at young individuals living in the regions with high unemployment rates.

JEL Classification: J61, H21, C35

Keywords: Migration, taxation, expected income, regional policy

* Corresponding author.

1 Introduction

During recent years migration in Finland has been characterised by a shift towards the areas of economic growth. If it continues this current trend will lead to the concentration of human capital in a few attractive regions (see e.g. Pekkala 2000; Ritsilä 2001).¹ At the same time, more distant regions are losing important future human capital, as the young and educated move away from them.

The role of interregional migration in the concentration of human capital and economic activities is a very important and problematic question from the perspective of regional policy. The key problems are the legitimacy of the different policy options and the effectiveness of the tools used in their implementation. The first question concerns how desirable in terms of efficiency and equality the trend towards concentration is (Lucas 1997, 758). The question of effectiveness is exceedingly important when new regional policies are being planned and implemented.

The question of individually focused policy measures is currently at the centre of public debate in Finland. For example, is it possible through income taxation policy to influence where people choose to live? In Norway different taxation schedules are applied in Northern Norway, including a reduction in personal income tax and an exemption from payroll tax for employers (Eikeland 1999). So far no such policies have been applied in Finland, but such measures have been seen as one way to combat out-migration. However, there has been no rigorous examination of whether and how migration decisions can be affected by individual taxation measures in the first place. In this respect we extend previous studies² and examine, for example, how significant an effect an increase in expected net income would have on migration decisions.

The empirical analysis of the paper is based on data from the Finnish Longitudinal Census File for period 1993–95. The data set used herein is a random sample of over 20,000

¹ The trend towards urbanisation is not unique. Many western and developing countries are experiencing similar problems (see e.g. Lucas 1997, 725; United Nations 1999).

² Previous studies on the impact of expected income gains on migration include Nakosteen and Zimmer (1980) for the United States, Robinson and Tomes (1982) and Islam and Choudhury (1990) for Canada and Haapanen (2000) for Finland. However, these studies do not consider taxation.

individuals. It includes information on personal and family status, economic activity and region of residence. The analysis focuses on persons of working age.

Our modelling is based on the human capital framework, where migration decisions are assumed among other factors to depend on expected net income. Estimations are implemented in two steps. Firstly, endogenous switching regression model of income determination is estimated and expected income after taxes is calculated for the alternatives of migration and non-migration. Secondly, the net income predictions are used as additional explanatory variables in a migration model.

Policies can however have different impacts on the propensity to move to peripheral regions compared to growth-centre regions. Therefore, in our migration model an individual can choose between three alternatives: (s)he can either (i) stay in the region of origin, (ii) migrate to a peripheral region, or (iii) migrate to a growth-centre region. In this framework it is possible to test whether the policy if implemented would have varying impacts on migration to the two different types of regions. This could not be done with the simple and commonly used model, in which an individual can only migrate or stay.

The discrete migration decisions are modelled with a random parameter logit model.³ In this flexible model the parameter of expected net income – and therefore the impact of expected net income on migration – can be specified as a function of set of explanatory variables, including a random factor. Thus the effect of expected net income on migration can vary among different types of individuals and regions of origin. This makes it possible to estimate which demographic groups and regions would be the most affected by a regional taxation policy. To be more precise, migration propensities are calculated before and after the implementation of taxation policies to examine their effectiveness. These calculations give guidelines on where to target the possible tax reductions. This in turn gives us information about the cost-effectiveness of the regional policy measures in question.

The remainder of the paper is organised as follows. Section 2 presents the theoretical background of the paper. Section 3 outlines the econometric methods, while the data and variables used in the empirical analysis are introduced in Section 4. The outcomes of the

³ Random parameters (“mixed”) logit models have been used in various applications e.g. in marketing and consumer research (Jain, Vilcassim, and Chintagunta 1994; Train 1998; Brownstone and Train 1999), but, to our knowledge, they have not been applied to migration problems.

estimations are reported in Section 5, and the policy implications are discussed in Section 6. Finally, Section 7 concludes the paper.

2 Theoretical framework

2.1 Human capital theory and migration

The theoretical setting of this paper is related to the human capital framework. This framework is based on the modelling work of Sjaastad (1962), Weiss (1971), Seater (1977) and Schaeffer (1985). In it migration is assumed to result from variations in individual economic utility in different locations. Furthermore, an individual is assumed to maximise his or her economic utility. Thus, relocation takes place if the expected economic utility gained from moving exceeds the economic utility achieved by staying in the present location. Heterogeneous individuals possess different utility functions, and consequently encounter differences in the net benefits of living in a specific location.

An important factor that affects economic utility, and hence the decision making of an individual, is her/his personal human capital reserve. Human capital can be considered as a heterogeneous asset, resulting from formal schooling, training and experience, etc. In addition, economic utility will depend on (expected) net income, which is a function of taxation. Formally, it is assumed that an individual i decides to migrate from location j to location k under the following utility maximisation process at a given time t :

$$E(R_{it}) = \max_{W, X} \left[E \int_0^T e^{-rt} \{U_{ik}(W_{ik}, X_{ik}) - U_{ij}(W_{ij}, X_{ij})\} dt - CM_{ijk} \right] \quad (1)$$

under the precondition

$$\int_0^T e^{-rt} (U_{ik} - U_{ij}) dt - CM_{ijk} \geq 0, \quad (2)$$

where $E(R_{it})$ is the net present value of the expected economic utility of that individual i . U_k is the expected utility level achieved in the alternative location k with assets X_{ik} and expected net income W_{ik} . Similarly, U_j is the expected utility achieved by living in the present location j with assets X_{ij} and expected net income W_{ij} . The direct costs involved in moving from location j to location k are contained in CM_{jk} . The expected utilities U_k and U_j , as well as the direct costs CM_{jk} , are formed as a result of personal, household and regional factors involved in the migration decision process. As a result of the rational decision making process, a decision in

favour of migration is reached when the expected utility gained from moving exceeds the direct costs of moving.

2.2 Taxation, migration and formation of income after taxes

Taxation in the region of origin and in the region of destination may have effects on migration decisions. Firstly, because it has a direct effect on expected net income and thus on the propensity to migrate. For example, Nakosteen and Zimmer (1980) examine the returns from interstate migration in the United States, while Robinson and Tomes (1982) and Islam and Choudhury (1990) compare the wages of movers with those of stayers in Canada. Haapanen (2000) considers the case of Finland. They all find that choosing whether to migrate depends positively on the expected gains in wages.

Secondly, taxation in the region of origin can have an indirect effect on migration, because the net wage rate influences labour supply decisions.⁴ According to theoretical models, a fall in taxation (an increase in the net wage rate) will cause an increase in labour supply, if the substitution effect dominates the income effect (i.e. if labour supply curve is upward sloping; see e.g. Connolly and Munro 1999). In addition, there has been considerable empirical evidence in the literature suggesting that employed people are less likely to move than non-employed people (see e.g. Ritsilä and Tervo 1999; Juarez 2000). Therefore, we would expect a reduction in the taxation of the region of origin to have a negative effect in the propensity to migrate.⁵

The implementation of a regional taxation policy presupposes knowledge about the formation of earned income after taxes so that the policy can be targeted at the right individuals and the costs can be assessed. Figure 1 illustrates how earned income was taxed in Finland in 1995.⁶ We can see that income taxation is progressive. For example, if a taxpayer earned 17,000 euro in 1995, his income tax rate was approximately 10%. If a taxpayer earned 35,000 euro, (s)he had to pay approximately 20% in taxes (twice as much).

⁴ It is unlikely that taxation in the region of destination would have an indirect effect on migration. In other words, it is unlikely that a change in the net wage rate in the other regions would significantly alter the amount an individual would be willing to work in the region of origin, everything else being equal.

⁵ Naturally, the impact of taxation may also depend on the form of taxation. For example, lump-sum, proportional and progressive taxes may each have different impacts on the labour supply and hence on the propensity to move.

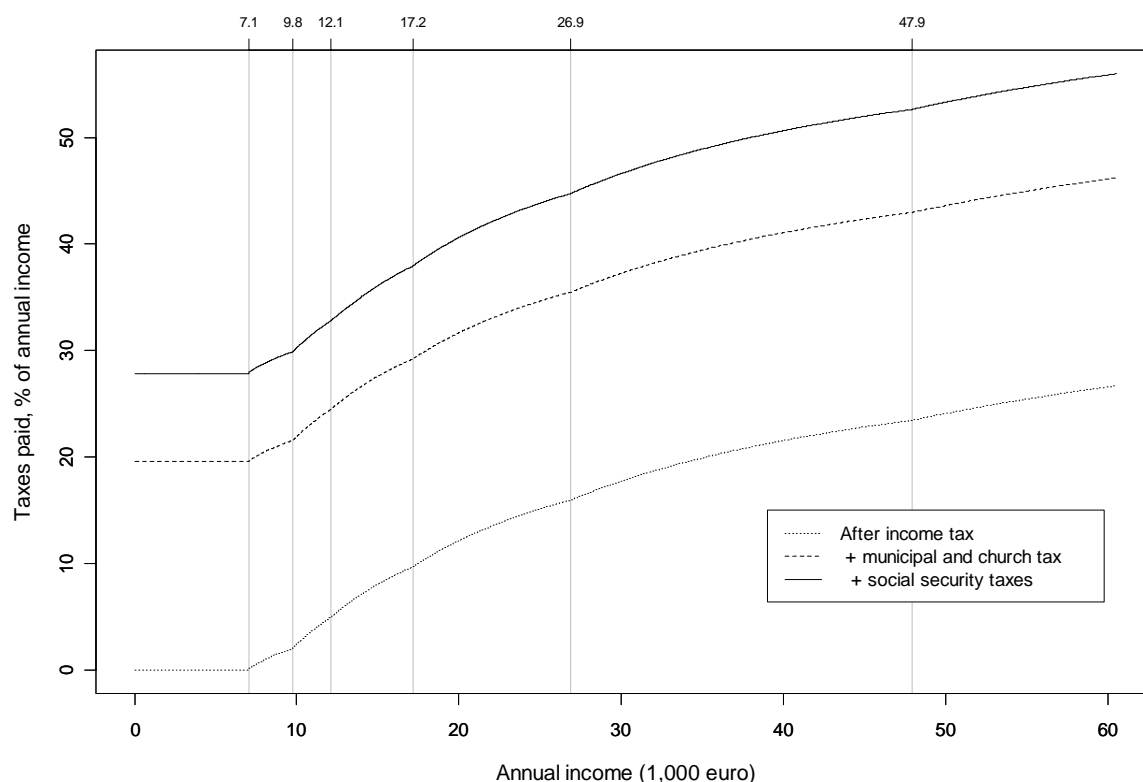


Figure 1. Illustration of progressive taxation in Finland in 1995⁷

Municipal and church taxes are levied at flat rates on taxable income. The municipal tax rate varied between 15% and 20% in 1995, depending on the municipality. Church tax is paid by the members of the Finnish Evangelical Lutheran (and Orthodox) churches at flat rates on taxable income. These rates varied between 1% and 2.25% in 1995.

Social security taxes consist of a sickness insurance premium, pension and unemployment insurance premiums. In 1995 a sickness insurance premium of 1.9% was collected on earned income (3.8% in the case of incomes exceeding 13,452 euro). Pension and unemployment insurance premiums were deducted at source at a rate of 6.4%. These facts on the Finnish taxation system were used when calculating income after taxes for 1995.

⁶ In addition, personal earned income subject to tax not only includes wages and salaries paid in money, but also directors' bonuses, commission, rental value of an employee's free housing, pensions and annuities, living and housing allowances, car benefits and unemployment benefits.

⁷ Mean values of municipal and church taxes were used in producing the figure. Tax deductions were not taken into account.

3 Empirical model specifications

In Section 3.1 we introduce the model that we used to predict the individual's net income after the migration period for both alternatives: staying in the region of origin and migrating. The income predictions are later used in the migration choice model presented in Section 3.2.

3.1 Modelling income predictions

Variations in income arise from differences in human capital, measured by education and work experience (see e.g. Mincer 1978). However, there may be other, unobserved factors which influence income determination and migration. Therefore, we proceed with a self-selection endogenous switching model (see e.g. Powers 1993; Maddala 1983; Poirier and Ruud 1981):

$$y_{ji} = \beta_j' x_{ji} + \varepsilon_{ji}, \quad j = 0, 1 \quad (3)$$

$$y_i = (1 - m_i) y_{0i} + m_i y_{1i}, \quad (4)$$

where y_{ji} is the annual net income of an individual i obtained by choosing alternative j and β_j is the parameter vector associated with income-determining attributes x_{ji} such as level of education and work experience. The observed income, y_i , which will be in state 1, if the premium ($y_{1i} - y_{0i}$) exceeds the cost of moving from state 0 (stay) to alternative state 1 (migrate), $m_i^* = \delta' z_i + u_i$. Otherwise, y_i is observed in state 0 and no movement of labour takes place. To be more specific, $m_i = 1$ if $\delta' z_i + u_i > 0$, $m_i = 0$ otherwise (an endogenous selection process). In the selection process, the difference between income offers will affect the worker's choice. Consequently, we combine the variables included in vector x_i with other personal characteristics to form a comprehensive vector z_i , which with parameter vector δ determines the sector process. Finally, ε_{ji} and u_i are disturbance terms.⁸

Thus, the complete switching regression model contains two income equations and a switching equation. Ordinary least squares (OLS) estimation of the above equations will provide unbiased estimates of income parameters only if u_i is uncorrelated with ε_{0i} and ε_{1i} .

⁸ The stochastic specification for the disturbance terms is $(\varepsilon_0, \varepsilon_1, u) \sim N(0, \Sigma)$. $\text{cov}(\varepsilon_0, \varepsilon_1)$ is not identified because each individual in the sample realises either y_0 or y_1 , but not both. Therefore, $\text{cov}(\varepsilon_0, \varepsilon_1)$ is set equal to zero. By means of normalisation $\text{var}(u)$ is set equal to one. If the estimated error covariances are non-zero, this provides evidence of non-random selection into one or both of the selection regimes (Powers 1993, 251).

For example, if unobserved preferences influence the selection process as well as income determination, then this assumption is violated and income predictions based on OLS are misleading. Therefore, the full information maximum likelihood (FIML) is used to jointly estimate the parameters.⁹ Then, expected incomes for the two migration alternatives are calculated as unconditional linear predictions

$$\hat{y}_{ji} = \hat{\beta}_j' x_{ji}, \quad j = 0, 1 \quad (5)$$

where $\hat{\cdot}$ denotes estimated values. These predictions are used in the migration model specified in the next section.

3.2 Modelling migration choices

Let us consider a migration model where an individual can choose between several alternatives instead of just two alternatives (migrate or stay). For the sake of generality, we assume that an individual i faces a choice among J mutually exclusive alternatives. It is also assumed that choices are made according to the well-known random utility maximisation hypothesis (see e.g. Ben-Akiva and Lerman 1985), where the (indirect) utility of alternative j is the sum of a deterministic component and a random component.¹⁰ The deterministic component of the utility depends on the attributes of the alternatives (such as expected income) and the characteristics of the individual (such as age and education). In our case, individual can choose between three alternatives: he or she can either (i) stay in the region of origin, (ii) migrate to a peripheral region, or (iii) migrate to a growth-centre region.

The traditional way to model choice situations has been to use the multinomial logit (MNL) model, because its likelihood function can be easily maximised due to a closed-form solution and because its results are convenient to interpret (Ben-Akiva and Lerman 1985). It is also well known that the choice probabilities of MNL have the IIA property. That is, for each individual, the ratio of the choice probabilities of any two alternatives is independent of the utility of any other alternatives. Hence, the tractability of the MNL model comes at a cost.¹¹

⁹ For a derivation of the likelihood function, see Maddala (1983, Ch. 9).

¹⁰ Although utility-based approaches to choice making have been popular, there are alternative ways of seeing the choice process in the literature. For example, Tversky's (1972) elimination-by-aspect decision making sees choice as a process involving the elimination of alternatives, with the process terminating when only a single alternative remains. However, it has several drawbacks, which have prevented its wider use (Pudney 1989, 122). See Ben-Akiva & Lerman (1985, 35–38) and Pudney for other examples.

¹¹ See Maddala (1995) a review of tests for the IIA property.

To obtain efficient and accurate estimates of the choice model parameters one should include a specification of the heterogeneity structure in the model: the presence of heterogeneity will alter cross-income elasticities, marginal rates of substitution between choices, and lead to IIA violations (Hensher 2001). We take this perspective in setting out a choice model that can incorporate the unobserved effects using random parameters logit model (RPL; see e.g. McFadden and Train 2000).

We derive the RPL model by assuming that the utility, $U(i, j)$, that individual i receives given a choice of alternative j is

$$U(i, j) = \alpha_{ji} + \theta_j' g_i + \phi_j' f_{ji} + \eta_{ji}' w_{ji} + \varepsilon_{ji}, \quad (6)$$

where α_{ji} is an alternative specific constant (fixed or random). θ_j is a vector of fixed coefficients (with a normalisation: $\theta_1 = \mathbf{0}$). ϕ_j is a vector of fixed coefficients. η_{ji} is a coefficient vector randomly distributed across individuals. g_i is a set of choice-invariant individual specific characteristics (such as age and education). f_{ji} and w_{ji} are vectors of alternative specific characteristics (such as expected income).

We first assume that the random parameters η follow a general distribution $g(\eta|\Omega)$. Estimation of the RPL model involves estimating the vectors Ω in addition to the fixed parameters as in the standard MNL model. In practice, we specify the random parameters as¹²

$$\eta_{ji} = \gamma_j + \delta_j' h_i + \sigma_j u_{ji}, \quad (7)$$

where $u_{ji} = N(0, 1)$ and $\eta_{ji} = N(\gamma_j + \delta_j' h_i, \sigma_j^2)$. h_i is the set of individual specific characteristics that influence the expected income parameter η . The unobserved disturbance term u_{ji} is individual and choice-specific.

For each individual, the migration choice probabilities will be

$$P(j|\eta) = \frac{\exp(U(i, j))}{\sum_{j=1}^J \exp(U(i, j))}, \quad i = 1, \dots, N; \quad j = 1, 2, \dots, J. \quad (8)$$

If the value of η was known for each individual, the solution to Equation (8) would be straightforward (resulting into the MNL model). However, η is unobserved, although it is

¹² A number of other error covariance structures can be specified in random parameters logit models (see e.g. Hensher 2001).

drawn from a known joint density function g . Thus, in order to obtain the unconditional choice probabilities for each individual the logit probability must be integrated over all values of η weighted by the density of η :

$$P(j) = \int_{\eta} \frac{\exp(U(i, j))}{\sum_{j=1}^J \exp(U(i, j))} g(\eta | \Omega) \partial \eta. \quad (9)$$

Examination of the above equation reveals that the choice probability is a mixture of MNL probabilities with g as the mixing distribution. The integrals in Equation (9) cannot be evaluated analytically since it does not have a closed-form solution. Therefore, the integrals in the choice probabilities are approximated using a Monte Carlo technique and the resulting simulated log-likelihood function is maximised.

4 Data

The impact of expected income on the migration decisions is tested with a one-percent random sample from the Finnish Longitudinal Census File. Statistics Finland has combined the population census with various employment registers maintained by the Ministry of Labour. The socioeconomic status of the sample individuals and their spouses is well documented: the data include information on personal and family status, labour market record, and regional characteristics of over 60,000 persons.¹³ The empirical analysis of this study mainly utilises data from the years 1993–1995.¹⁴

The observed trends in the spatial distribution of population and employment are fundamentally related to three key choices made by the working population: (i) where to live, (ii) whether or not to work, and (iii) where to work. The model used here focuses exclusively on the workplace and residential decisions and does not consider the labour market participation decision. Furthermore, because we are interested in the effects of tax reductions, only individuals aged between 16 and 65 with a positive annual income in 1995 were included in the sample.¹⁵ Self-employed, retired and foreign-born individuals were excluded

¹³ The data do not allow us to use households as the unit of analysis, because we do not know what individuals belong to the same households. However, we do have wide range of household variables, which should adequately control for dependencies in migration decision making.

¹⁴ During the period under study, 1993–95, the Finnish economy was recovering from recession, the unemployment rate remaining exceptionally high. The speed of migration was also considerable lower compared to situation at present (see e.g. Pekkala 2000, 18).

¹⁵ An annual income greater than 1,000 euro was considered positive.

from the sample, as their income determination is likely to differ from that of the rest of the population. After these restrictions and omitting observations with missing information we are left with 20,213 observations. Figure 2 illustrates their labour market status during the last week of 1993. Most of them were employed (73.0%) or unemployed (13.9%), but some (3.2%) had no earned income at all in 1993.

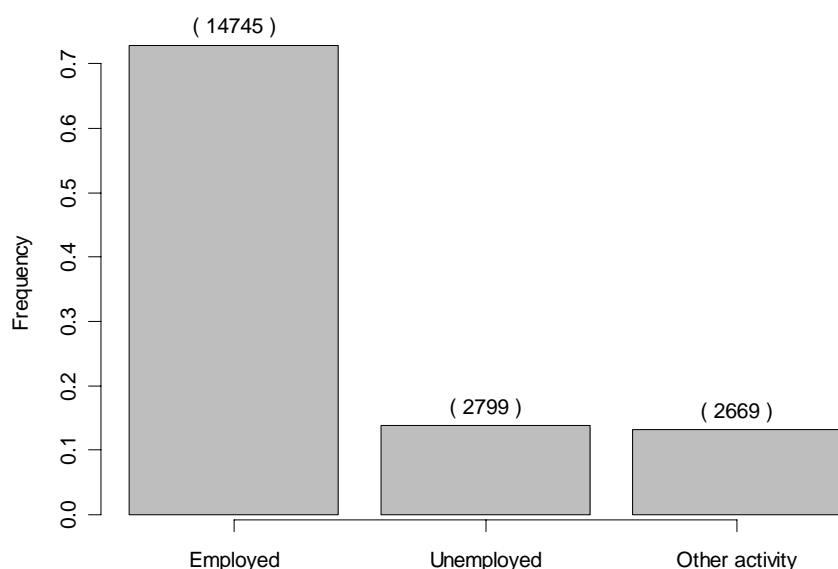


Figure 2. Main type of activity during the last week of 1993 (number of observations in brackets)

The present analysis exploits a classification of growth-centre regions and peripheral regions (Figure 3). This classification is formed using information on the net migration rates and population figures for the destination subregions in 1995: a region is classified as a growth-centre region, if it has a positive net in-migration rate and its population is larger than 50,000 inhabitants.¹⁶ The growth-centre regions are Helsinki, Porvoo, Salo, Tampere, Turku, Vaasa, Jyväskylä, Kuopio and Oulu, and are also characterised by high wage levels and low unemployment rates.¹⁷ The other 75 subregions are mostly peripheral and stagnating regions, although, in Finnish terms, they include some of the bigger towns (Lappeenranta, Rovaniemi).

¹⁶ The regional division will not alter if the population is kept between 44,000 and 60,000.

¹⁷ See e.g. Statistics Finland's Kuntafakta (statistics on Finnish municipalities).

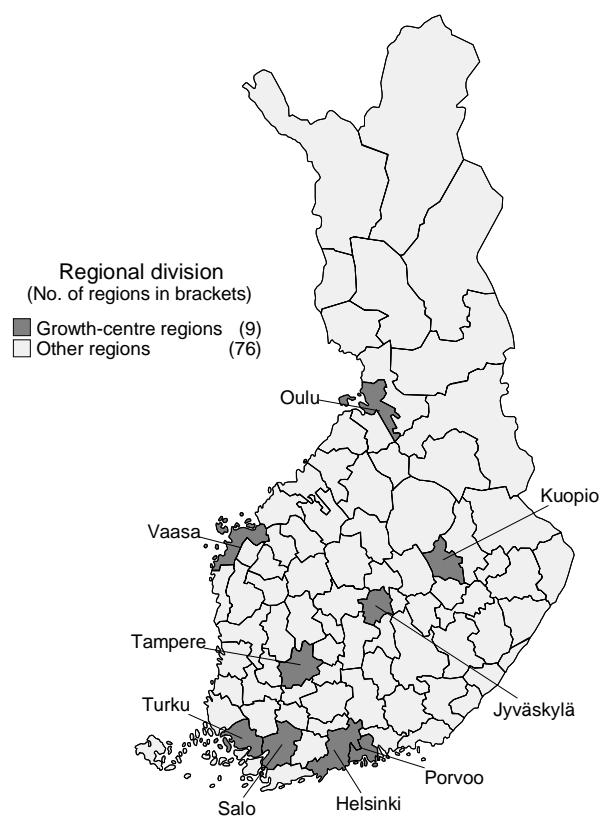


Figure 3. Regional division into growth-centre and other regions

The sample data will allow us to determine whether and where an individual migrated in 1994 by using information on the subregion of residence and the above regional division.¹⁸ An individual is classified as a growth-centre migrant, if (s)he migrates in 1994 and his/her region of destination is one of the growth-centre regions. In the sample of 20,213 individuals, 558 individuals (2.76%) migrated in 1994, of whom 285 individuals migrated to growth-centre regions (1.35%) and 273 individuals migrated to other regions (1.41%).

Measure of income after taxes (net income) is constructed using the information on the Finnish tax system described in Section 2.2. The state, municipal, church and social security taxes are subtracted from the individual's earned annual income.¹⁹ We know in which subregion an individual lives, but not in which municipality; therefore, we have to use a proxy for the municipal tax paid by each individual. It was calculated as a population-weighted average of all the municipal taxes in his/her subregion.

¹⁸ We used data from 84 subregions (NUTS4, "seutukunta" in Finnish). The subregion of Åland was excluded, as it has many distinctive characteristics (e.g. self-regulation, isolated geographical location and a Swedish-speaking majority).

Next we briefly introduce the variables that are used to explain individuals' migration decisions and income determination. Mean values and variable descriptions are given in the Appendix (Table A1). All the explanatory variables, except for net expected income, are measured at the region of origin before the migration period in the migration and income equations (see Figure 4).

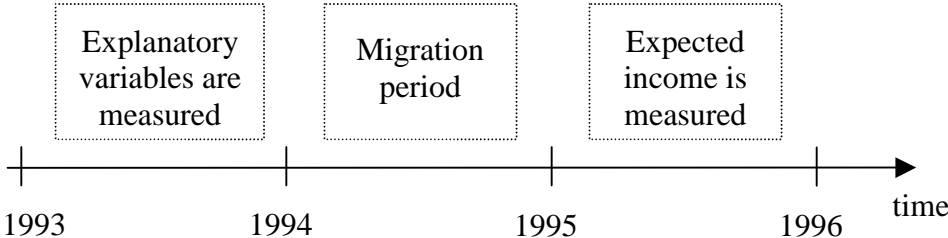


Figure 4. Composition of the data

The explanatory variables include age, sex, and level of education (academic degrees). Information on spouses is used to construct dummies that indicate whether the individual is married or cohabiting and whether his/her spouse is employed. Because the presence of children can influence migration decisions, we have created two dummies that reveal whether an individual has school-aged children, under school-aged children only, or does not have children at all. Work experience is defined simply as number of months of employment during 1987–93. We also use information about the home-ownership and commuting status, the number of months of employment, and whether the individual was employed during the last week of 1993.

We use several variables that describe the individual's region of origin. They include the regional rate of unemployment and the degree of urbanisation in the region of origin. In addition, the individual's distance to the closest growth-centre region is used as a measure of geographical location. It was calculated using distances by road (10 km). Firstly, the distance between two regions was calculated by using a distance matrix based on municipalities. Then in each subregion the most populated municipality was chosen to represent the location of the region.

¹⁹ For simplicity, tax deductions were not taken into account when income after taxes was calculated.

5 Estimation results

5.1 Switching regression for expected income

In this section we report the estimation results of the switching regression model. The model contains two components: two income equations and a switching equation. The entire model is estimated by full information maximum likelihood. The first set of dependent variables are net income for the migration and staying alternatives in 1995. They are specified with explanatory variables measured in 1993. We have followed the standard convention of specifying the income equations with education, experience, and experience squared as the central explanatory variables. In addition to these human capital variables, we have included some individual characteristics like sex, age, and age squared. Furthermore, we have included information on the degree of urbanisation in the income equation for stayers, because it reflects their opportunities in the regions of origin.²⁰

The switching equation is specified with variables commonly used in the migration literature, such as age, education, home-ownership, and status of spouse and children. As speculated earlier, the difference in income between the migration and staying alternatives may well play a role in the selection process, affecting the willingness to migrate. Therefore, the variables that determine the individual's net income are also included in the selection equation. In addition, information on the region of origin is used.

The estimations results are reported in Table 1. The results fully correspond to our prior expectations. Because they are not the focus of this paper, we shall omit further discussion of them; they are merely used to generate expected net income for the migration and staying alternatives. However, it is worth reporting the results of the covariance parameters, since their statistical significance implies that selection to migration categories is not random.

²⁰ The variable is not relevant for migrants, because it is measured in the region of origin. For them, it does not reflect their income opportunities, because their region of residence has changed.

Table 1. Parameter estimates of the switching regression model

Variable	Income equations				Switching equation	
	Stayer		Migrant			
Constant	-0.168**	(0.056)	-0.009	(0.524)	-2.024**	(0.259)
Sex (male = 1)	0.195**	(0.010)	0.193**	(0.064)	-0.106**	(0.035)
Age	0.597**	(0.032)	0.715**	(0.267)	0.108	(0.155)
Age squared	-0.071**	(0.004)	-0.094*	(0.037)	-0.036	(0.022)
Lower academic degree	0.412**	(0.016)	0.469**	(0.114)	0.330**	(0.061)
Higher academic degree	0.694**	(0.021)	0.803**	(0.150)	0.407**	(0.068)
Spouse employed					-0.164**	(0.041)
School-aged children					-0.194**	(0.050)
Homeowner					-0.031	(0.040)
Living in region of birth					-0.204**	(0.036)
Work experience	0.068**	(0.007)	0.103**	(0.017)	0.119**	(0.027)
Work experience squared	0.005**	(0.001)			-0.004	(0.003)
Commuting					0.419**	(0.040)
Degree of urbanisation	0.049**	(0.003)			-0.013	(0.014)
Distance to the closest growth-centre					0.010**	(0.002)
<i>Covariance parameters</i>						
$\sigma_{00} = 0.663^{**} (0.003)$, $\sigma_{u0} = 0.932^{**} (0.005)$, $\sigma_{11} = 0.725^{**} (0.026)$, $\sigma_{u1} = 0.014 (0.208)$						

Notes: First the estimated parameter is given, followed by the asymptotic standard errors in brackets. Log-likelihood: -21,535.29. Number of observations $N = 20,213$. The staying alternative is the reference state in the switching equation. All variables measured at the region of origin (see Appendix, Table A1). * (**) = statistically significant at the 5% (1%) level.

The terms σ_{uj} , $j = 0, 1$, measure the correlation of the income equation disturbance term in migration alternative j with that of the switching regression. The term σ_{u0} is highly significant and shows that the migration choice equation has a strong correlation with the income equation in the staying alternative. Hence, it is likely that the OLS estimates would have been seriously biased, had they been used.²¹ The sign of σ_{u0} conform to intuition: someone expecting, for some unobserved reason, a high income offer in the present region is also more likely to stay in that region.

Note finally the intuitive result that the variance of the migration income offers (σ_{11}) is larger than that of the staying income offers (σ_{00}), reflecting the uncertainty of income determination following migration or the smaller number of observations in the migration category than in the staying category.

²¹ Having said that, the OLS estimates were not very different from the estimates obtained from the switching regression model. The results are not reported here but are available from the corresponding author upon request.

5.2 Random parameters logit model for migration decision

The net income predictions from the switching regression model are used as additional explanatory variables in a random parameter logit model of migration choice. In the migration model an individual can choose between three mutually exclusive alternatives: (s)he can either (i) stay in the region of origin, (ii) migrate to a peripheral region, or (iii) migrate to a growth-centre region. The first alternative is used as a normalising base category.

In order to determine whether any parameters can be eliminated, we estimated the random parameters model with various parameter restrictions. Starting from the most general model we searched for the most parsimonious specification: group-by-group equality and zero restrictions were imposed and tested. The final estimation results are reported in Table 2; see also Equations (6) and (7).

Table 2. Parameter estimates of the random parameters logit model

<i>Non-random parameters</i>	Alternative			
	Staying (reference state)	Migrating to periphery	Migrating to growth- centre	
Constant		-1.089* (0.523)	-8.407** (0.650)	
Age		-0.339** (0.108)	-0.527** (0.128)	
Academic degree		0.209 (0.149)	0.209 (0.149)	
Spouse is employed		-0.415** (0.151)	-0.595** (0.191)	
Under school-aged children only		0.003 (0.184)	-0.355 (0.210)	
School-aged children		-0.436* (0.180)	-1.246** (0.253)	
Homeowner		-0.803** (0.149)	0.460** (0.177)	
Months of employment		-0.028 (0.019)	0.012 (0.023)	
Commuting		0.553** (0.178)	1.250** (0.187)	
Regional rate of unemployment		0.015 (0.018)	0.073** (0.025)	
Living in region of birth		-0.831** (0.142)	-1.073** (0.163)	
Living in growth-centre region		-2.904** (0.243)	5.591** (0.212)	
Distance to the closest growth-centre		-0.095** (0.011)	0.112** (0.011)	
Initial earnings		-0.017 (0.092)	-0.158 (0.119)	
<i>Constant of the random parameter</i> ln(predicted annual income)	<i>Parameters influencing</i> <i>Heterogeneity in mean</i>			<i>Derived s.d. of the</i> <i>parameter</i> <i>distribution, σ</i>
	Age	Homeowner	Regional rate of unemployment	
14.941**	-1.356**	3.335**	-0.294*	0.018
(3.221)	(0.498)	(1.002)	(0.117)	(9.665)

Notes: First the estimated parameter is given, followed by the asymptotic standard errors in brackets. See Appendix (Table A1) for the variable definitions and mean values. Log-likelihood: -2,111.12. Number of observations $N = 20,213$. Replications for simulated probabilities: 500. * (**) = statistically significant at the 5% (1%) level.

We tried other specifications for the random parameter of $\ln(\text{predicted annual income})$, but variables age, homeowner, and regional rate of unemployment seemed to describe accurately the variation in the income expectations parameter of the utility functions. This is reflected on the insignificance of the derived standard deviation of the parameter distribution, σ . We also estimated the model with different expected income parameters for each utility function (j), but the parameters were very similar to each other and thus the equality restriction was not rejected. A comparison of the random parameters logit model against the multinomial logit model²² (log-likelihood of $-2,124.65$) indicates an important contribution to the overall statistical fit from a less restrictive specification of the unobserved effects.

Before considering the impact of expected net income on the migration choices more closely, we briefly report the results for the other significant variables. Because the age-variable affects the utilities not only directly but also indirectly through the expected income parameter, our interpretations are based on simulated probabilities. As expected, old people are less likely to migrate than young. The older the migrant, the fewer will be the years of payoff from the human capital investment in migration, while the cost of migration remains just as high, which helps to explain why migration diminishes with age. In addition, younger individuals are expected to have lower psychological costs, because of fewer local social ties.

As Mincer (1978) points out, family considerations appear to influence migration decisions.²³ People are tied to other people and hence the individual's propensity to move is decreased in the presence of children. The results presented in Table 2 show that in the presence of school-aged children, a parent becomes less willing to move and is even less likely to move to a growth-centre region. One can also be tied to a house, as might be the case for a homeowner. Our results indicate that owning a house reduces the probability to move to peripheral regions (see Table 4).

The results for the distance variable are illustrated in Table 3. On the basis of simulated probabilities, we find that individuals living in the growth-centre regions are less likely to stay in their present region and more likely to migrate to another growth-centre region than individuals living in the peripheral regions. For individuals living in peripheral regions, an increase in the distance to the closest growth-centre region reduces (marginally increases) the

²² The MNL results are available from the corresponding author upon request.

²³ In general, we did not find gender differences in migration, which is not surprising given that we have controlled for various household characteristics.

probability of moving to other peripheral regions (growth-centre regions). Living in one's region of birth reduces willingness to move. The larger the unemployment rate in the region of origin, the more likely one is to move to a growth-centre region. The unemployment rate can be viewed as a proxy for labour market diversification and career opportunities in the region of origin.

Table 3. Simulated choice probabilities by distance to the closest growth-centre region

	Alternative		
	Staying	Migrating to periphery	Migrating to growth-centre
Individual 1	0.9594	0.0026	0.0380
Individual 2	0.9684	0.0313	0.0003
Individual 3	0.9794	0.0201	0.0005
Individual 4	0.9863	0.0127	0.0009

Notes: Individual 1 lives in a growth-centre region and thus his/her distance to the closest growth-centre region is 0 km. Individual 2 (3, 4) lives in a peripheral region and his/her distance to the closest growth-centre region is 50 (100, 150) km. The probabilities are simulated using estimates given in Table 2. They are calculated as averages over all observations. Replications for simulated probabilities: 1,000.

Finally, we consider the impact of expected net income on the migration decision. Table 2 suggests that the impact of expected net income on the migration decision is dependent on age, home-ownership and the regional rate of unemployment, but again simulated probabilities are necessary to get a clearer picture of the migration decision process.

Table 4 reports the simulated probabilities before and after a 10% increase in expected net income for the *staying* alternative.²⁴ Hence, this policy would increase individuals' income in their region of origin. Absolute and percentage changes in the probabilities are also reported. The simulations were performed with different values for the variables age, homeowner and regional rate of unemployment. The reported figures are averages over all observations. On average, the policy would increase individuals' probability of staying in the region of origin by 2.09% and it would decrease their probability of moving to a growth-centre (peripheral) region by 73.61% (73.54%).

²⁴ Of course, we can also calculate the probabilities before and after a 10% increase in expected net income for all migration alternatives. For brevity, the results are not reported but are available from the corresponding author upon request.

Table 4. Simulated probabilities before and after a 10% increase in expected net income for the staying alternative

Alternative	Average individual	Age		Homeowner		Regional rate of unemployment	
		20 years	40 years	No	Yes	15%	25%
<i>Before increase in income</i>							
Staying	0.9724	0.9634	0.9811	0.9689	0.9754	0.9783	0.9671
Migrating to periphery	0.0135	0.0177	0.0097	0.0182	0.0085	0.0124	0.0141
Migrating to growth-centre	0.0141	0.0189	0.0092	0.0129	0.0161	0.0093	0.0188
<i>After increase in income</i>							
Staying	0.9927	0.9934	0.9931	0.9896	0.9966	0.9973	0.9896
Migrating to periphery	0.0036	0.0029	0.0035	0.0062	0.0009	0.0014	0.0043
Migrating to growth-centre	0.0037	0.0037	0.0034	0.0042	0.0025	0.0013	0.0061
<i>Absolute change in probability</i>							
Staying	0.0203	0.0300	0.0120	0.0207	0.0212	0.0190	0.0225
Migrating to periphery	-0.0099	-0.0149	-0.0061	-0.0120	-0.0076	-0.0110	-0.0098
Migrating to growth-centre	-0.0104	-0.0152	-0.0059	-0.0087	-0.0136	-0.0080	-0.0127
<i>Percentage change in probability</i>							
Staying	2.09%	3.12%	1.22%	2.13%	2.17%	1.94%	2.33%
Migrating to periphery	-73.61%	-83.73%	-63.41%	-65.95%	-89.02%	-88.68%	-69.47%
Migrating to growth-centre	-73.54%	-80.25%	-63.55%	-67.31%	-84.60%	-86.36%	-67.64%

Notes: The probabilities are simulated using estimates given in Table 2. They are calculated as averages over all observations. Replications for simulated probabilities: 1,000. The average predicted annual net income in the staying alternative is 8,674 euro.

The simulated probabilities yield the expected result that the younger one is, the larger will the effect of expected income changes on the probabilities of staying and moving. If the individual is 20 years old (40 years old), his/her probability of staying would on average increase by 3.12% (1.22%) and his/her probability of migration would on average fall by some 82% (63%). Similar conclusions can be drawn from the absolute changes in the probabilities. For example, for individuals aged 20, a 10% increase in net income in the region of origin would increase their average probability of staying by 0.03.

The elasticity of staying with respect to an increase in expected net income in the region of origin hardly changes with home-ownership. For homeowners, the probability of staying increases by 2.17%, and for those who do not own their home by 2.13%. However, the absolute probability of moving to a growth-centre (periphery) falls considerably more (less) for homeowners than for those who do not own their home. Note, however, that their initial migration probabilities are different. The elasticity of migration is considerably greater with home-ownership.

The elasticity of staying is greater in regions of high unemployment (2.33%) than in regions of low unemployment (1.94%). In absolute figures, the average probability of moving to a growth-centre region is reduced more for individuals living in regions of high unemployment than for individuals living in regions of low unemployment. This is a good outcome, if it is desired to retain more people in the peripheral regions.

6 Discussion of policy implications

Figure 5 shows percentage net migration into the Finnish subregions in 1999. We can see a geographical shift of population towards the growth-centre regions. In fact, the subregional concentration of population has strengthened in the recent past. Net migration into the growth-centre and peripheral regions was 8,459 and -5,194 individuals, respectively, in 1995 as against 17,374 and -14,596 individuals in 1999. Could a reduction in the personal taxation of the biggest net migration loser regions be an answer to the problem (see the eleven regions coloured white in Figure 5)?

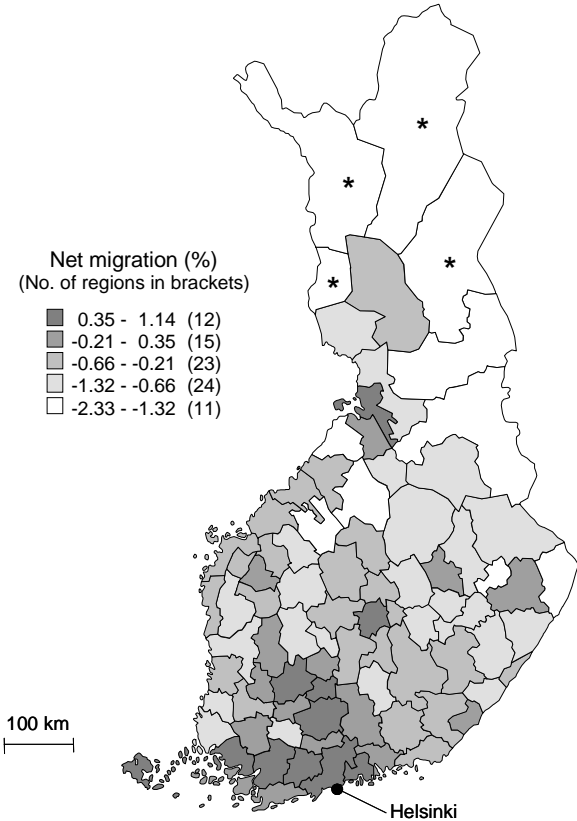


Figure 5. Net migration in 1999, % (Source: Statistics Finland’s Kuntafakta), * indicates regions used in the policy analysis

Firstly, it would mean that fewer people would move from the lowered taxation regions and thus reduce out-migration. Secondly, it would probably increase in-migration to those regions. Therefore, the selected regions would lose fewer inhabitants. However, it is difficult to evaluate the aggregate affects of such policies. That is, how many more people would stay in the area and how many more people would move into an area of reduced taxation. We shall focus on the first problem.

To approach the problem, we have to decide how this tax reduction is implemented in practice. Because income tax is progressive in Finland, it may not be desirable to use it: for example, only individuals with earned income in excess of 11,000 euro would be affected by the policy in 2001. Instead, social security taxes and municipal tax seem more eligible candidates because they are proportional to earned income and are paid by everyone with earned income. In the latter case, however, the government would have to compensate the municipalities for these lost tax revenues so as not to throw their already poor economies even further out of balance.

Here we consider the implications of two regional taxation policies targeted at individuals living in the four northernmost subregions (see Figure 5). In the first policy their annual net income is increased by 10%, for example with the help of social security taxes. In the second policy their annual net income is increased by 25%. This latter net growth in income can be achieved, for example, by assuming that the individuals who are affected by the policy do not have to pay municipal tax.²⁵

Table 5 illustrates the cost-effectiveness of such policies for 1999. We can see that there were 26,898 employed individuals with an average annual net income of approximately 9,137 euro in the four subregions. Therefore, the two policies would cost approximately 24.585 and 61.435 million euro to the government, respectively. Statistics also show that 3,284 (1,927) individuals migrated from (into) the four subregions in 1999, resulting in a net in-migration deficit of -1,357. Our simulations, which rely on the fact that the unemployment rate is approximately 25% in those regions, imply that on the average the out-migration probability would diminish by some 74% and 84% as a result of the first and second policies,

²⁵ Because the average annual income, municipal tax and taxes paid in total were approx. 13,453 euro, 18% and 32%, respectively, in those subregions in 1999, the policy would increase their annual net income by some 27%.

respectively.²⁶ Since approximately one third of the migrants are employed and are thus directly affected by the policy, the out-migration would fall to 2,474 and 2,365 individuals, respectively. Of course, the policy would also attract other than employed persons to stay in these and increase migration into them. Therefore, we estimate that the net out-migration would be significantly more balanced after both taxation policies. However, the 10% increase in annual net income appears to be a more sensible alternative than the 25% increase, since the latter, more expensive policy, does not reduce the gap much more than the cheaper one.

Table 5. Costs and effects of two taxation policies

	Before the taxation policy ⁽¹⁾	After 10% increase in annual net income	After 25% increase in annual net income
<i>Costs</i>			
Employed individuals	26,898	(2)	(2)
Annual income per individual, euro	13,453	(2)	(2)
Annual net income per individual, euro	9,137	10,051	11,421
Annual costs per individual, euro	-	914	2,284
Total annual costs, euro	-	24,584,772	61,435,032
<i>Effects</i>			
In-migration	1,927	(2)	(2)
Out-migration	3,284	2,474 ⁽³⁾	2,365 ⁽³⁾
Net in-migration	-1,442	-547	-438

Notes: The four subregions used in the illustration are given in Figure 5. ¹⁾ Figures are given for 1999 (*Sources:* Finnish Tax Administration & Statistics Finland's Kuntafakta). ²⁾ Figure is assumed to be unchanged after the policy. ³⁾ Figures are calculated using estimates in Table 2.

7 Conclusion

In this paper, we analyse migration choices in Finland focusing on whether migration decisions can be affected by income taxation policies. Our findings have important implications for public policy. First, we show that the decision to migrate is affected by net income expectations, as the human capital theory predicts. Second, we find that the size of the impact of expected net income on migration depends on age, home-ownership and regional unemployment rate. In addition, we consider the effectiveness of income taxation in reducing out-migration from stagnating regions. Our simulations imply that the effectiveness of such policy would be increased, if the reduction in taxation could be targeted at young individuals living in regions with high unemployment rates. Hence, we provide important evidence on where to target possible tax reductions.

²⁶ The simulated percentage changes are averages over all sample individuals living in the peripheral regions.

Acknowledgements

This paper is part of Mika Haapanen's doctoral studies in the Finnish Postgraduate Programme in Economics (FPPE). Both authors would like to thank the Yrjö Jahnsson Foundation for financial support. An earlier version of this paper was presented at FPPE's research seminar in Labour Economics in Helsinki. We appreciate the valuable comments by Arnaud Chevalier and other participants. We would also like to thank Hannu Tervo for his helpful comments and suggestions on this paper.

Appendix

Table A1. Variable definitions and descriptive statistics

Variable	Scale	Operational definition	Mean
Sex	(0, 1)	1 = male; 0 = female	0.497
Age	Continuous	Age in years divided by 10	3.641
Age squared	Continuous	(Age/10) squared	14.427
Lower academic degree	(0, 1)	1 = lower academic degree; 0 = otherwise	0.097
Higher academic degree	(0, 1)	1 = higher academic degree; 0 = otherwise	0.069
Academic degree	(0, 1)	1 = academic degree; 0 = otherwise	0.166
Spouse employed	(0, 1)	1 = married or cohabiting and spouse's main activity is employment during the last week of 1993; 0 = otherwise	0.465
Under school-aged children only	(0, 1)	1 = under school-aged children only; 0 = otherwise	0.119
School-aged children	(0, 1)	1 = school-aged children; 0 = otherwise	0.327
Homeowner	(0, 1)	1 = homeowner; 0 = otherwise	0.422
Initial earnings	Continuous	Annual earnings in 1993 (10,000 euro) = annual income from labour plus self-employment income and work-related transfers, such as unemployment insurance and sick pay.	1.647
Months of employment	(0-12)	Months of employment in 1993	8.520
Work experience	Continuous	No. of months of employment during 1987–93 divided by 10	6.083
Work experience squared	Continuous	Work experience squared	44.469
Commuting	(0, 1)	1 = commuting from the municipality in the last week of 1993; 0 = otherwise	0.336
Living in region of birth	(0, 1)	1 = living in region of birth; 0 = otherwise	0.539
Living in growth-centre region	(0, 1)	1 = living in growth-centre region, i.e. in the region of Helsinki, Salo, Porvoo, Tampere, Turku, Vaasa, Jyväskylä, Kuopio or Oulu; 0 = otherwise	0.476
Degree of urbanisation	(0-10)	Proportion of population living in built-up areas	7.749
Distance to the closest growth-centre	Continuous	Individual's distance to the closest growth-centre using distances by road (10 km)	5.679
Regional rate of unemployment	Continuous	Regional rate of unemployment, %	21.997
Ln(predicted annual income)	Continuous	Natural logarithm of predicted annual income for 1995 (1,000 euro). Calculated using the endogenous switching model (see Table 1).	2.091

Notes: All variables measured at the region of origin in 1993 if not otherwise stated.

References

- Ben-Akiva, M. and Lerman, S. R. (1985) *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT Press, Cambridge.
- Brownstone, D. and Train, K. (1999) Forecasting new product penetration with flexible substitution patterns, *Journal of Econometrics*, 89, 109–129.
- Connolly, S. and Munro, A. (1999) *Economics of the Public Sector*, Prentice Hall Europe, London.
- Daganzo, C. (1979) *Multinomial Probit: The Theory and Its Application to Demand Forecasting*, Academic Press, London.
- Eikeland, S. (1999) Personal incentives in Norwegian regional policy, Nordregio working paper 5/1999.
- Haapanen, M. (2000) Expected earnings and interregional migration, University of Jyväskylä, School of Business and Economics, Working paper No. 221.
- Hensher, D. A. (2001) The sensitivity of the valuation of travel time savings to the specification of unobserved effects, *Transportation Research E*, 37, 129–142.
- Islam, M. N. and Choudhury, S. A. (1990) Self-selection and intermunicipal migration in Canada, *Regional Science and Urban Economics*, 20, 459–472.
- Jain, D. C., Vilcassim, N. J. and Chintagunta, P. K. (1994) A random-coefficients logit brand-choice model applied to panel data, *Journal of Business and Economic Statistics*, 12, 317–328.
- Juarez, J. P. (2000) Analysis of interregional labor migration in Spain using gross flows, *Journal of Regional Science*, 40, 377–399.
- Lucas, R. E. B (1997) Internal migration in developing countries, in: Rosenzweig, M. R. and Stark, O. (eds.) *Handbook of Population and Family Economics*, Volume 1B, 721–798.
- Maddala, G. S. (1983) *Limited-dependent and Qualitative Variables in Econometrics*, Cambridge University Press, Cambridge.
- Maddala, G. S. (1995) Specification tests in limited dependent variable models, in: Maddala, G. S., Phillips, Peter C. B. and Srinivasan, T. N. (eds.) *Advances in Econometrics and Quantitative Economics: Essays in Honor of Professor C. R. Rao* (Blackwell, Oxford), 1–49.
- McFadden, D. and Train, K. (2000) Mixed MNL models for discrete response, *Applied Econometrics*, 15, 447–476.
- Mincer, J. (1978) Family migration decisions, *Journal of Political Economy* 86, 749–773.

- Nakosteen, R. A. and Zimmer, M. (1980) Migration and income: the question of self-selection, *Southern Economic Journal*, 46, 840–851.
- Pekkala, S. (2000) Regional convergence and migration in Finland 1960–90, Doctoral Dissertation, Jyväskylä Studies in Business and Economics, University of Jyväskylä.
- Poirier, D. J. and Ruud, P. A. (1981) On the appropriateness of endogenous switching, *Journal of Econometrics*, 16, 249–256.
- Powers, D. A. (1993) Endogenous switching regression models with limited dependent variables, *Sociological Methods & Research*, 22, 248–273.
- Pudney, S. (1989) *Modelling Individual Choice: The Econometrics of Corners, Kinks and Holes*, Blackwell, Cambridge.
- Ritsilä, J. (2001) Studies on spatial concentration of human capital, Doctoral Dissertation, Jyväskylä Studies in Business and Economics, University of Jyväskylä.
- Ritsilä, J. and Tervo, H. (1999) Regional Differences in the Role of Migration in Labour Market Adjustment: The Case of Finland, in: Crampton, G. (ed.) *Regional Unemployment, Job Matching and Migration*, Series on European Research in Regional Science (Pion, London), 166–182.
- Robinson, C. and Tomes, N. (1982) Self-selection and interprovincial migration in Canada, *Canadian Journal of Economics*, 15, 474–502.
- Schaeffer, P. (1985) Human capital accumulation and job mobility, *Journal of Regional Science*, 25, 103–114.
- Seater, J. J. (1977) A unified model of consumption, labour supply, and job search, *Journal of Economic Theory*, 14, 349–372.
- Sjaastad, L. A. (1962) The costs and returns of human migration, *Journal of Political Economy*, 70 (Supplement), 80–93.
- Stern, S. (1997) Simulation-based estimation, *Journal of Economic Literature*, 35, 2006–2039.
- Train, K. (1998) Recreation demand models with taste differences over people, *Land Economics*, 74, 230–239.
- Tversky, A. (1972) Elimination by aspects: a theory of choice, *Psychological Review*, 79, 281–299.
- United Nations (1999) World urbanization prospects: the 1999 revision, United Nations Population Division.
- Weiss, Y. (1971) Learning by doing and occupational specialization, *Journal of Economic Theory*, 4, 189–198.