On the Dynamics of Interstate Migration: Migration Costs and Self-Selection

Christian Bayer,*
Falko Juessen†‡
Universität Dortmund

First version: February 15, 2006
This version: October 16, 2006

Abstract

This paper develops a tractable dynamic microeconomic model of migration decisions that is aggregated to describe the behavior of interregional migration. Our structural approach allows us to deal with dynamic self-selection problems that arise from the endogeneity of location choice and the persistency of migration incentives. Keeping track of the distribution of migration incentives over time has important consequences, because the dynamics of this distribution influences the estimation of structural parameters, such as migration costs. For US interstate migration, we obtain a cost estimate of somewhat less than one-half of an average annual household income. This is substantially less than the migration costs estimated by previous studies. We attribute this difference to the treatment of the dynamic self-selection problem.

KEYWORDS: Dynamic self-selection, migration, indirect inference
JEL-codes: C61, C20, J61, R23

*University of Dortmund, Faculty for Economics and Social Sciences, Department of Economics, 44221 Dortmund; phone: +49-231-755-3246; fax: +49-231-755-3027; email: christian.bayer@wiso.uni-dortmund.de

†University of Dortmund, Faculty for Economics and Social Sciences, Department of Economics, 44221 Dortmund; phone: +49-231-755-3291; fax: +49-231-755-3069; email: f.juessen@wiso.uni-dortmund.de

‡We would like to thank conference participants at the North American Summer Meeting of the Econometric Society 2006, the Society of Economic Dynamics Meeting 2006, and the Annual Meeting of the European Economic Association 2006 for their helpful comments and suggestions. Part of this paper was written while C. Bayer was visiting fellow at Yale University and Jean Monnet fellow at the European University Institute. He is grateful for the support of these institutions. Financial support by the Rudolf Chaudoire Foundation is gratefully acknowledged. The Research has been supported by DFG under Sonderforschungsbereich 475. We would like to thank Christian Wogatzke for excellent research assistance. A previous version of the paper has been circulated under the title "A generalized options approach to aggregate migration with an application to US federal states".

1 Introduction

Migration decisions are important economic decisions. Migration allows individual agents to smooth their income and is an important way of adjustment to macroeconomic shocks (Blanchard and Katz, 1992). Many factors influence the decision to migrate and there is a vast empirical literature that links migration decisions to economic incentives (see Greenwood, 1975, 1985, and 1997 and Cushing and Poot, 2004 for survey articles). At the same time, most of this literature has remained relatively silent about the actual costs of migration to individual agents. Nevertheless, migration costs are surely a structural parameter of high interest (Sjaastad, 1962).

So, there has recently been a small number of studies that actually do report estimates on migration costs. Davies, Greenwood, and Li (2001) report a cost estimate of about US$ 180,000 for each migration between US states, and Kennan and Walker (2003, 2006) conclude that, all other things equal, migration costs are between US$ 176,000 and US$ 270,000.[1] In terms of average annual income, this magnitude of migration costs corresponds to roughly 4-6 average annual household incomes. At any rate, such an estimate appears very high and even the authors of these studies are somewhat sceptical about their findings.

Kennan and Walker (2003) suggest that some kind of omitted variable problem may drive the high cost-estimate. In particular, they suggest that an unobservable wage component is correlated to the decision to stay. We argue that the endogeneity of the location choice will always lead to such correlation. In fact, this paper's first result is that it is necessary to keep track of the unobservable distribution of migration incentives over time to obtain an unbiased estimate of migration costs.

This motivates us to develop a tractable microeconomic model of migration which can be aggregated and used to describe the simultaneous evolution of migration incentives and migration rates at an aggregate level. Our model picks up the general idea that migration can be understood as an investment into human capital (Sjaastad, 1962). In particular, the migration-decision problem is closely related to the decision problem for discrete investment projects or lumpy investment.

For the lumpy investment setup, Caballero and Engel (1999) develop a methodological framework that allows them to estimate micro-level investment costs from aggregate data only. We extend their work to migration decisions. This means that we first develop a structural model of the representative microeconomic problem of migration for

[1] These estimates do not yet include mark-ups for distance and other factors that influence the psychic costs of migration. Return migration is usually associated with lower, but still substantial costs.

heterogeneous households and in a second step, this model is used to derive the evolution of the distribution of migration incentives. This evolution of incentives determines aggregate migration in turn.

We simulate this model and estimate migration costs via indirect inference (see Gourieroux, Monfort, and Renault, 1993 and Smith, 1993). Particularly, we apply Gourieroux, Monfort, and Renault's (1993) method of simulated moments. We estimate migration costs to be about US$ 20,000, which is somewhat less than one-half of the average annual income. This cost estimate is substantially lower than the cost estimates reported by previous studies. Moreover, we show that applying the techniques used in other papers, we would obtain higher cost estimates also from data generated by a simulation of our structural model. Consequently, we conclude that keeping track of the distribution of migration incentives over time has an important influence on the estimation of migration costs. This finding extends the role of self-selection problems to a dynamic setup, which so far have been highlighted in static frameworks (see for example Borjas, 1987, Borjas, Bronars, and Trejo, 1992, Tunali, 2000, and Hunt and Mueller, 2004).

Finding more reasonable cost estimates parallels the results of the investment literature, in which much more reasonable estimates of adjustment costs were obtained when fixed adjustment costs to capital were included into dynamic models. For migration, the issue of fixed and sunk costs was emphasized in the real-options approach by Burda (1993) and Burda et al. (1998). However, these papers only look at migration as a once and for all decision, so that they preclude return migration. Moreover, the papers do not study the evolution of migration incentives, to which past migration decisions feed back.

Taking into account these feedbacks, we extend the structural approaches of Davies, Greenwood, and Li (2001) and Kennan and Walker (2006) and suggest a fully structural model of migration that is based upon dynamic optimization and hence takes into account the dynamic character of the migration decision. This allows us to track the dynamic evolution of migration incentives at the macroeconomic level, but it comes at the cost that we have to reduce the model to a bi-regional setup for numerical feasibility. One distinct feature of our model is that it allows us to infer the structural microeconomic parameters of the migration decision from aggregate data only.

The remainder of this paper is organized as follows: Section 2 gives a brief discussion of the difficulties of estimating structural migration models when the population dynamically self-selects into its preferred region. The section develops the main motive of our paper and illustrates why migration costs are hard to identify by standard (dis-

crete choice) estimation techniques. Thereafter, Section 3 presents a tractable dynamic microeconomic model of the migration decision which assumes that an agent maximizes future expected well-being by location choice. In Section 4, we show how to aggregate the model. We derive the contemporaneous law of motion of the distribution of migration incentives and aggregate migration rates, taking into account heterogeneity at the microeconomic level. We provide the results of a numerical simulation analysis in Section 5 to give an idea of how the proposed model actually behaves. Section 6 finally confronts the model with aggregate data on migration between US states and presents the estimates of the structural parameters of the model, particularly the estimates of migration costs. Section 7 concludes and an appendix provides detailed proofs as well as details on the data employed.

2 What makes migration costs so hard to identify?

Most micro studies and now also more macro studies on migration link the individual migration decision to a probabilistic model in which agents migrate if the gain in utility terms obtained by migration,

$$\left(u_{it}^{\text{move}} - u_{it}^{\text{stay}}\right) = \gamma x_{it} + \nu_{it}, \tag{1}$$

is large enough and exceeds some threshold value $\bar{c}$.[2] This threshold value $\bar{c}$ can be interpreted as migration costs in utility terms. The vector of covariates x_{it} is composed of information that describes the economic incentives to migrate, i.e. the gains from migration.

For example, x_{it} could contain data on remuneration, on labor market conditions, and on amenities for both the home and the destination region. The vector of parameters γ measures the sensitivity of the migration decision to these economic incentives. The stochastic component ν_{it} reflects differences across agents, omitted migration incentives, and/or some variability of migration costs.

Typically, we are interested in the structural parameters γ and $\bar{c}$ and hence would estimate some version of (1) to infer these parameters. Unfortunately, such direct approach is very difficult due to the unobservability of the *potential* migration gains to the outside observer. To illustrate this point, suppose an agent only cares about the difference in income between home and destination region.

In such setting, x_{it} were simply a measure of relative income potentials for an agent

[2] See for example Davies, Greenwood, and Li (2001), Hunt and Mueller (2004) or Kennan and Walker (2006).

which she can realize by location choice. A rational agent then moves to the region where she earns the most, provided that her migration costs are covered by the discounted present value of the differences in future incomes.

However, the econometrician can only observe the income that an agent realizes in the region in which she is currently living. Therefore, the other, the unobserved, potential income has to be proxied. Typically, it is proxied by an income a similar agent realizes in the other region.[3] At a macro level, this often means replacing agent-specific income differences by average income differences across regions, see for example Davies, Greenwood, and Li (2001).

If we proxy the unobservable income difference x_{it} for individual i in equation (1) by the average income difference $\bar{x}_{.t}$ between source and destination region, then we obtain

$$\left(u_{it}^{\text{move}} - u_{it}^{\text{stay}}\right) = \gamma \bar{x}_{.t} + \underbrace{\gamma\left(x_{it} - \bar{x}_{.t}\right) + \nu_{it}}_{\text{composed error term}}. \tag{2}$$

The composed error term $\gamma\left(x_{it} - \bar{x}_{.t}\right) + \nu_{it}$ now also includes the idiosyncratic component of income differences $\eta_{it} := \left(x_{it} - \bar{x}_{.t}\right)$. Since we do not want to base our following argument on a classical measurement error or omitted variable problem, we may assume that the idiosyncratic component to the income difference η_{it} is orthogonal to the average income difference.[4] For the ease of exposition, also suppose that the agent really just cares about income, so that the true stochastic component is actually identical to zero, $\nu_{it} \equiv 0$.

Under these assumptions, we can rewrite (2) as

$$\left(u_{it}^{\text{move}} - u_{it}^{\text{stay}}\right) = \gamma \bar{x}_{.t} + \gamma \eta_{it}. \tag{3}$$

In this equation, the regression residual only captures the distribution of idiosyncratic potential income differences around the mean.

While the migration decision is deterministic to the individual in this setting, it is stochastic to the econometrician due to his lack of knowledge of η_{it}. If the econometrician were to know the distribution of the unobserved component η_{it}, he could nonetheless estimate γ with a suitable probabilistic discrete choice model. However, assuming one

[3] One example is the paper of Hunt and Mueller (2004) that does a Mincer-type wage regression to obtain the unobservable income potential. A similar example can be found in Burda et al. (1998) or Kennan and Walker (2006).

[4] Alternatively, one could think of η_{it} as being the unexplained residual of a Mincer-type wage regression and $\bar{x}_{.t}$ being the income component that is explained by all observable characteristics of the agent. Our line of argument applies to this microeconomic interpretation, too.

of the standard distributions for η_{it}, e.g. a logistic distribution, is problematic.

Suppose agents are heterogeneous with respect to potential incomes, so that the idiosyncratic component η_{it} has a non-degenerated distribution. In particular, assume that η_{it} is initially normally distributed as displayed in Figure 1 (a), so that in the initial situation a probit model were appropriate. The figure displays the distribution of migration incentives, i.e. potential incomes, $x_{it} = \bar{x}_{.t} + \eta_{it}$. Low values of this sum imply that income in region A is favorable, high values of this sum imply better income prospects in region B. Correspondingly, all agents with $\bar{x}_{.t} + \eta_{it} < 0$ decide to live in region A and they decide to live in region B otherwise if we assume zero migration costs for the moment. In other words, the agents self-select into the region that is favorable for them.[5]

As a result, the distribution of income differences changes for the next period. No agent who lives in region A prefers to live in region B. This means that for those agents who live in region A the distribution of income differences is as displayed in Figure 1 (b). Effectively, the right-hand part of the distribution in Figure 1 (a) has been cut, because all agents with higher income in region B have actually chosen B as the region to live in.

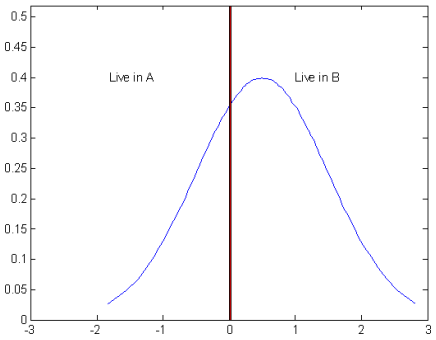
It can be seen that the migration incentives $\bar{x}_{.t}+\eta_{it}$ are no longer normally distributed conditional on a household living in region A. Since the estimation residual $\gamma\eta_{it}$ in our setup results from a linear transformation of the migration incentive $\bar{x}_{.t} + \eta_{it}$, also the estimation residual $\gamma\eta_{it}$ is no longer normally distributed. Accordingly, the distributional assumptions to estimate (1) by standard maximum likelihood techniques are no longer fulfilled.

Even adding a normally distributed idiosyncratic income shock does not reestablish a normal distribution of income differences if income differences are sufficiently persistent. Figure 1 (c) displays how mild idiosyncratic shocks alter the distribution displayed in Figure 1 (b) . Again, the distribution is different from the standard distributions assumed in the estimation of discrete-choice models. The colored-in region indicates the set of agents that will migrate from A to B after the idiosyncratic shocks.

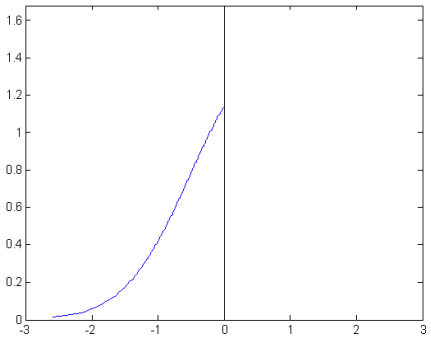
Besides idiosyncratic shocks, also aggregate shocks to the income difference $\bar{x}_{.t}$ influence the migration decisions of agents. Figure 1 (d) shows the distribution of migration incentives as in Figure 1 (c) but after an adverse shock to region A. By comparing Figures 1 (c) and 1 (d), one can see that the shape of the distribution after migration differs

[5] This self-selection is driven directly by the heterogeneity of the agents with respect to potential incomes, but is does not reflect immanent and fixed differences of the regions as in Borjas (1987) and Borjas, Bronars, and Trejo (1992).

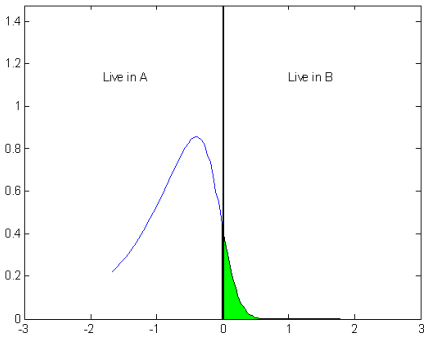
Figure 1: Distribution of potential income in region B relative to A



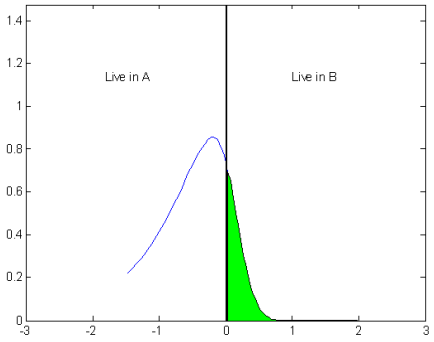
(a) overall population



(b) conditional on living in region A after migration



(c) conditional on living in region A after migration and idiosyncratic shocks



(d) conditional on living in region A after migration, idiosyncratic, and aggregate shocks

(the not colored-in region). In consequence, the distribution of migration incentives will not be strictly stationary, it will evolve over time, and it will depend on the history of aggregate shocks.

Hence, the distribution differs in two important aspects from those assumed in standard discrete-choice models. Firstly, it will not be one of the standard distributions considered. Secondly, it will display a dynamic behavior as a result of aggregate shocks.

Now, how does this correspond to an unreasonable estimate of migration costs? If $\bar{c}$ is normalized to 1, the parameter γ has a straightforward interpretation. It measures the sensitivity of migration decisions to income incentives and its inverse $\frac{1}{\gamma}$ is exactly the income differential at which an average agent is just indifferent between moving and not moving. Or to put it differently, $\frac{\bar{c}}{\gamma}$ is the money measure of average migration costs.

In turn, this implies that a bias in γ directly translates into a bias in estimated migration costs. And with the distribution of migration incentives misspecified, γ will be estimated with a bias most probably. The misspecification of the distribution of migration incentives has two aspects. One is that the distribution will always be non-standard, i.e. neither normal nor logistic. The second aspect is that the distribution also changes over time as a result of aggregate shocks to income and the triggered migration decisions.

To put it simply: agents are in a certain region most likely because they are better off living there. Because of this self-selection, the distribution of unobserved migration incentives is most likely not symmetric (see Greenwood, 1985, pp. 533). Additionally, it displays a dynamic behavior. Accordingly, one needs to keep track of the evolution of the incentive distribution and standard techniques to deal with self-selection cannot be applied in a straightforward way. Therefore, we develop a model based on dynamic optimal migration decisions in the presence of persistent shocks to income. This model can then be aggregated and used to simulate the evolution of migration and its incentives over time.

3 A simple stochastic model of migration decisions

We consider an economy with two regions, A and B. For simplicity, this economy is assumed to be inhabited by a continuum of infinitely lived agents of measure 1. We model the economy in discrete time and at each point in time an agent has to decide in which region to live and work. First, we consider the decision problem of an individual agent. For simplicity, we drop the index i that has denoted the specific individual before, but use this index to indicate regions, $i = A, B$.

Living in region i at time t gives the agent utility $\tilde{w}_{it}$. Although $\tilde{w}_{it}$ is a catch-all

variable for migration incentives, which can be interpreted as wage income, employment prospects, amenities, utility from social networks etc., we refer to $\tilde{w}_{it}$ as income for simplicity.

The agent discounts future utility by factor $\beta < 1$ and maximizes the discounted sum of expected future utility by location choice. Moving from one region to the other is not costless to an agent. When an agent moves, she is subject to a disutility c_t that enters additively in her utility function.

Hence, the instantaneous utility function $u(i, j, t)$ is given by

$$u\left(i, j, t\right) = \tilde{w}_{it} - \mathbb{I}_{j \neq i} c_t \tag{4}$$

for an agent that has lived in region j before and now lives in region i. Here, $\mathbb{I}$ denotes an indicator function, which equals 1 if the agent has moved from region j to i and 0 if the agent already lived in region i before.

Both variables, migration incentive (income $\tilde{w}_{it}$) and moving costs (c_t), are stochastic in our model. They vary over time and across individuals, but are observed by the agent before she chooses her location. The agent knows the distribution of both components of her utility function and forms rational expectations about future incomes and migration costs.

Since migration costs are stochastic and hence vary, not all individual agents who face the same income differential will actually take the same migration decision. In this sense, the individuals in our model are heterogenous and to the outside observer the migration decision is stochastic.

With both $\tilde{w}_{it}$ and c_t being stochastic, the potential migrant waits not only for good income opportunities but also for low migration costs. In her migration decision, she thus takes into account two option values. One is the value to wait and learn more about future incomes and the other is to wait and search for lower migration costs.

Migration costs themselves depend on many factors and may include both physical and psychic costs of migration (Sjaastad, 1962), but the factors that determine migration costs are not constant. For example, search costs to find a new job and accommodation evolve with market conditions, the disutility of living separated from a family or spouse changes over time, just as marital status itself is neither constant nor irreversible. We pick up the variability in migration costs c_t by assuming them to be independently and identically distributed according to a distribution function G.

The distribution of migration incentives, $\tilde{w}_{it}$, is assumed to be log-normal. In particular, we assume that log income, w_{it}, follows an AR(1) process with normally distributed

innovations ξ_{it} and autoregressive coefficient ρ :

$$\ln(\tilde{w}_{it}) =: w_{it} = \mu_i(1-\rho) + \rho w_{it-1} + \xi_{it}. \tag{5}$$

This process holds for the whole continuum of agents and each agent draws her own series of innovations ξ_{it} for both regions. The expected value of log income in region i is μ_i. The innovations ξ_{it} are composed of aggregate as well as idiosyncratic components. They have mean zero, are serially uncorrelated, but may be correlated across regions A,B (see Section 4.2).

Income and cost distributions, together with the utility function and the discount factor define the decision problem for the potential migrant. This is an optimization problem, which is described by the following Bellman equation:

$$V(j, c_t, w_{At}, w_{Bt}) = \max_{i=A,B} \left\{ \exp(w_{it}) - \mathbb{I}_{\{i \neq j\}} c_t + \beta \mathbb{E}_t V(i, c_{t+1}, w_{A,t+1}, w_{B,t+1}) \right\}. \tag{6}$$

In this equation, $\mathbb{E}_t$ denotes the expectations operator with respect to information available at time t.[6]

The optimal policy is relatively simple. The agent migrates from region j to region i if and only if the costs of migration are lower than the sum of the expected value gain $\beta \mathbb{E}_t [V(i, c_{t+1}, w_{A,t+1}, w_{B,t+1}) - V(j, c_{t+1}, w_{A,t+1}, w_{B,t+1})]$ and the direct benefits of migration $\exp w_{it} - \exp w_{jt}$. This means that the agent migrates if and only if

$$c_t \leq \exp w_{it} - \exp w_{jt} + \beta \mathbb{E}_t [V(i, c_{t+1}, w_{A,t+1}, w_{B,t+1}) - V(j, c_{t+1}, w_{A,t+1}, w_{B,t+1})]. \tag{7}$$

The expected value difference

$$\mathbb{E}_t [V(i, c_{t+1}, w_{A,t+1}, w_{B,t+1}) - V(j, c_{t+1}, w_{A,t+1}, w_{B,t+1})]$$

may for example reflect different income expectations. Holding income expectations constant, the difference of the expected values also reflects the differences in expected future migration costs.

Since the costs of migration, c_t, are assumed to be i.i.d., expected costs at time $t+1$ do not depend on information available at time t. Moreover, the distribution of

[6] For technical reasons, we need to assume boundedness of ξ_{it}, so that ξ_{it} is in fact only *approximately* normal. The bounds to ξ_{it} turn the optimization problem into a bounded returns problem, which is easier to solve. But the bounds to ξ_{it} can be chosen arbitrarily wide (but finite) so that the distribution of $\tilde{w}_{it}$ approximates the log-normal distribution arbitrarily close. Existence and uniqueness of the value function is proved in the appendix.

future incomes $(w_{A,t+1}, w_{B,t+1})$ is a function of only (w_{At}, w_{Bt}), because w_{it} follows a Markov-process. This allows us to summarize the expected value difference by a function $\Delta V(w_{At}, w_{Bt})$ of only (w_{At}, w_{Bt}), which is defined as

$$\Delta V(w_{At}, w_{Bt}) := \beta \mathbb{E}_t \left[V(B, c_{t+1}, w_{A,t+1}, w_{B,t+1}) - V(A, c_{t+1}, w_{A,t+1}, w_{Bt,+1}) \right]. \quad (8)$$

Substituting (8) for the value difference in (7) gives a critical level of costs $\bar{c}$ at which an agent living in region A is just indifferent between moving and not moving to region B. This threshold is

$$\bar{c}(w_A, w_B) := \exp w_B - \exp w_A + \Delta V(w_{At}, w_{Bt}). \quad (9)$$

To put it differently, a person moves from A to B if and only if

$$c_t \leq \bar{c}_A := \bar{c}(w_{At}, w_{Bt}).$$

Conversely, a person living in region B moves to region A if and only if

$$c_t \ \leq \bar{c}_B := -\bar{c}(w_{At}, w_{Bt}).$$

Note that $\bar{c}$ can be positive as well as negative. If $\bar{c}$ is positive, region B is more attractive. If it is negative, region A is more attractive and a person living in region A would only have an incentive to move to region B if migration costs were negative.

4 Aggregate migration and the dynamics of income distributions

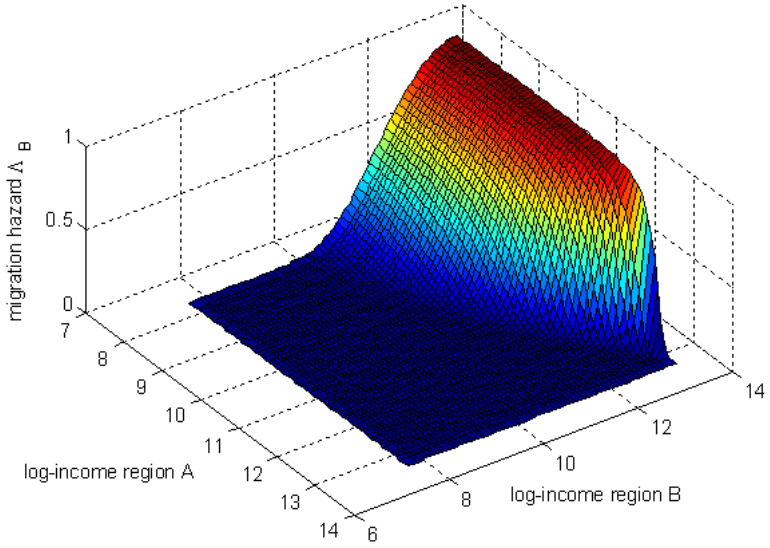
4.1 Aggregate migration

Given this trigger rationale for migration, the hazard rate

$$\Lambda_i(w_A, w_B) := G(\bar{c}_i(w_A, w_B)) \ , \ i = A, B$$

is the probability that a person in region i moves to the other region if she faces the potential incomes (w_A, w_B). This means that the likelihood of a person to move equals the probability that her migration costs realize below the threshold value $\bar{c}_i$. Since we assumed a continuum of agents, the actual fraction of migrating agents with income pair (w_A, w_B) is equal to this hazard rate, too. Figure 2 displays an example of a microeconomic migration-hazard function that stems from the optimization problem (6). The figure shows how different income combinations change the probability to migrate

Figure 2: Hazard-rates for migration from region A to region B conditional on potential incomes



from region A to B.

Now, consider the distribution F_t of (potential) incomes (w_A, w_B) and household locations. Suppose this income distribution is the distribution after the income shocks ξ_{it} have been realized, but before migration decisions have been taken. Let f_{it} denote the conditional density of this income distribution, conditional on the household living in region i at time t. Then, the actual fraction $\bar{\Lambda}_{it}$ of households living in i that migrate to the other region evaluates as

$$\bar{\Lambda}_{it} := \int \Lambda_i (w_A, w_B) \cdot f_{it} (w_A, w_B) \, dw_A dw_B. \tag{10}$$

This means that the aggregate migration hazard, $\bar{\Lambda}_{it}$, is a convolution of the microeconomic adjustment hazard Λ_i and the conditional income distribution f_{it}. In other words, the aggregate migration hazard can be thought of as a weighted mean of all microeconomic migration hazards, weighted by the density of income pairs (w_A, w_B) from distribution F_t.

4.2 Dynamics of income distributions

The distribution F_t itself (and hence f_{it}) evolves over time and is a result of direct shocks to income just as it is a result of past migration. We need to characterize the law of

motion for F_t to close our model and to obtain the sequence of aggregate migration rates.

4.2.1 The effect of migration on income distributions

Recall that the distribution F_t is the joint distribution of potential incomes and household locations. In order to follow the evolution of F_t we thus need to characterize both the evolution of the fraction P_{it} of households living in each region as well as the conditional distribution of incomes f_{it} (conditional on a household actually living in a specific region i).

The proportion of households living in region i at time $t+1$ is a result of migration decisions at time t. The law of motion for P_{it} is given by

$$P_{it+1} = \left(1 - \bar{\Lambda}_{it}\right) P_{it} + \bar{\Lambda}_{-it} P_{-it}. \tag{11}$$

The first part of the sum reflects the fraction of households that remain in region i, where $\left(1 - \bar{\Lambda}_{it}\right)$ is the probability to stay in region i. The second part is the fraction of households that migrate from region $-i$ to region i.

Since the microeconomic migration hazard depends on (w_A, w_B), different potential incomes in both regions result in different propensities to migrate. In consequence, migration changes not only the fraction P_{it} of households living in region i at time t, but also the conditional distribution of income, f_{it}. For example, households living in region A, earning a low current income, w_A, but facing a substantially higher potential income in B, w_B, are very likely to migrate. As a result, the number of those households strongly decreases after migration decisions have been taken, while the number of households facing a smaller income differential changes less.

These considerations form the backbone of our argument. The distribution of migration incentives is a result of past migration decisions, and we can express the new density of households with income (w_A, w_B) in region i after migration, $\hat{f}_{it}$, by

$$\hat{f}_{it}(w_A, w_B) = \left[1 - \Lambda_{it}(w_A, w_B)\right] \frac{f_{it}(w_A, w_B) P_{it}}{P_{it+1}} + \Lambda_{-it}(w_A, w_B) \frac{f_{-it}(w_A, w_B) P_{-it}}{P_{it+1}}. \tag{12}$$

The first product and part of the sum gives the fraction of households that remain in region i. In this product, the probability $\left[1 - \Lambda_{it}(w_A, w_B)\right]$ is again the probability to stay in region i. The term $f_{it}(w_A, w_B) P_{it}$ weights this probability and is the *un*conditional income density for region i before migration has taken place. To obtain again the conditional density, the unconditional income density, $f_{it}(w_A, w_B) P_{it}$, is divided by P_{it+1},

which is the fraction (or probability) of households living in region i after migration (i.e. in time $t+1$).

Analogously, the second part of the sum is constructed: $\Lambda_{-it}(w_A, w_B)$ is the probability to migrate from the other region, $-i$, to destination region i, $f_{-it}(w_A, w_B) P_{-it}$ is the unconditional income density for region $-i$, and dividing by P_{it+1} conditions for living in region i after migration.

4.2.2 The effect of income shocks on the income distribution

Besides migration, also shocks to income change the distribution of income pairs, F_t. These shocks can be purely idiosyncratic or may effect all individuals in the economy. For a single agent, we can decompose the total shock ξ_{it} to her potential income in region i (see equation 5) into an aggregate component θ_{it} and an individual-specific component ω_{it} :

$$\xi_{it} = \theta_{it} + \omega_{it}, \quad i = A, B.$$

The aggregate shock θ_{it} for region i hits all agents equally and changes their potential income for region i. Note that this shock does not depend on the actual region the agent is living in. For example, a positive shock $\theta_{At} > 0$ increases the *potential* income in region A for agents that are currently living in this region as well as for agents that are currently living in region B. They realize this potential income by deciding to actually live in region A. The correlation ψ_θ between θ_A and θ_B measures the importance of the economy-wide business cycles relative to the size of region-specific aggregate fluctuations.

However, aggregate shocks are typically only a minor source of income variation for an agent. Agents differ in various personal characteristics that result in different income profiles over time. Individuals differ in their skills and while the demand may grow for the skill of one person, demand may deteriorate for another person's skills. This heterogeneity is captured by the idiosyncratic shocks $(\omega_{At}, \omega_{Bt})$. If ω_{At} is positive, income prospects of the individual agent increase in region A. The correlation ψ_ω between ω_A and ω_B reflects economy-wide demand shifts for a person's individual skills.

Since we assume aggregate and idiosyncratic shocks to be independent, the variance of the total shock to income, ξ_{it}, is the sum of the variances of idiosyncratic and aggregate shocks: $\sigma_\xi^2 = \sigma_\omega^2 + \sigma_\theta^2$.

Persistency in incomes is captured by the autoregressive parameter ρ in equation (5). We abstain from the inclusion of permanently fixed individual differences (fixed effects) primarily because this makes the model numerically much more tractable.[7]

[7] If we were to include fixed effects that reflect different types of agents, the model had to be solved for each different type of agent just as it is now solved for the single type of agent.

Idiosyncratic shocks, aggregate shocks, and the persistency of the income process determine the transition of the distribution of income incentives after migration to the distribution of migration incentives before migration in the next period. The income distribution at the beginning of period $t+1$, F_{t+1}, results from adding idiosyncratic and aggregate shocks to the distribution of income after migration in period t, $\hat{F}_t$, of which $\hat{f}_{it}(w_A, w_B)$ is the conditional density, see (12). When a household has income w_{it+1} in period $t+1$, this can result from any possible combination of w_{it} and $\xi_{it+1} = \theta_{it+1} + \omega_{it+1}$ for which

$$w_{it+1} = \mu_i(1-\rho) + \rho w_{it} + \theta_{it+1} + \omega_{it+1} \tag{13}$$

holds. Solving this equation for w_{it} we obtain

$$w_i^*(w_{it+1}, \theta_{it+1}, \omega_{it+1}) := w_{it} = \frac{w_{it+1} - (\theta_{it+1} + \omega_{it+1})}{\rho} - \mu_i \frac{(1-\rho)}{\rho}. \tag{14}$$

This $w_i^*(w_{it+1}, \theta_{it+1}, \omega_{it+1})$ is the time-t potential income in region i that is consistent with a future potential income of w_{it+1} and realizations of shocks $\theta_{it+1} + \omega_{it+1}$ at the beginning of period $t+1$. Now suppose that both kinds of shocks, θ and ω, have been realized. Then, $w_{A,B}^*$ is a one-to-one mapping of future income (w_{At+1}, w_{Bt+1}) to current income (w_{At}, w_{Bt}).

The conditional density of observing the future income pair (w_{At+1}, w_{Bt+1}) can thus be obtained from a retrospective. The income pair (w_A^*, w_B^*) of past incomes corresponds uniquely to a future income pair (w_{At+1}, w_{Bt+1}). Consequently, we can express the density of the income distribution at time $t+1$ using the income distribution after migration $\hat{F}_t$, and its conditional density $\hat{f}_{it}$. The density of the income distribution F_{t+1} conditional on the region and the vector of shocks is given by

$$\begin{aligned} &f_{it+1}(w_A, w_B | \theta_{At+1}, \theta_{Bt+1}, \omega_{At+1}, \omega_{Bt+1}) \\ &\qquad = \hat{f}_{it}(w_A^*(w_A, \theta_{At+1}, \omega_{At+1}), w_B^*(w_B, \theta_{Bt+1}, \omega_{Bt+1})). \end{aligned} \tag{15}$$

Weighting this density with the density of the idiosyncratic shocks $h(\omega_{At+1}, \omega_{Bt+1})$ yields the density of observing the future income pair (w_A^*, w_B^*) together with the idiosyncratic shock $(\omega_{At+1}, \omega_{Bt+1})$:

$$\hat{f}_{it}(w_A^*(w_A, \theta_{At+1}, \omega_{At+1}), w_B^*(w_B, \theta_{Bt+1}, \omega_{Bt+1})) \cdot h(\omega_{At+1}, \omega_{Bt+1}).$$

Integrating over all possible idiosyncratic shocks $(\omega_{At+1}, \omega_{Bt+1})$ gives the density f_{it+1} of the income distribution before migration in period $t+1$ for a certain combination

of aggregate shocks $(\theta_{At+1}, \theta_{Bt+1})$:

$$f_{it+1}(w_A, w_B|\theta_{At+1}, \theta_{Bt+1}) = \int \hat{f}_{it}\left(w_A^*(w_A, \theta_{At+1}, \omega_A), w_B^*(w_B, \theta_{Bt+1}, \omega_B)\right) \cdot h(\omega_A, \omega_B)\, d\omega_A d\omega_B. \quad (16)$$

For given aggregate shocks, this new distribution determines migration from region i to region $-i$ according to equation (10) for time $t+1$.

The evolution of income distributions can thus be summarized as follows. Between two consecutive periods, the conditional distribution of potential incomes first evolves as a result of migration decisions, moving the density from f_{it} to $\hat{f}_{it}$. Thereafter, the distribution is again altered by aggregate and idiosyncratic shocks to income, moving the density from $\hat{f}_{it}$ to f_{it+1}. The latter density now determines migration decisions in time $t+1$, starting the cycle over again. In other words, migration incentives are not only a result of past income shocks, but also a result of past migration decisions. Keeping track of the distributional dynamics of migration incentives is at the heart of our model. This is the difference to most other empirical models of migration.

5 Simulation analysis

5.1 Numerical aspects

The first step in solving the model numerically is to obtain a solution to (6). We do so by value-function iteration.[8] For this value-function iteration, we first approximate the bivariate process of potential incomes for an individual agent in regions A and B

$$\begin{pmatrix} w_{At} \\ w_{Bt} \end{pmatrix} = w_t = \mu(1-\rho) + \rho w_{t-1} + \xi_t \quad (17)$$

by a Markov chain.[9] Because w_A and w_B are correlated through the correlation structure in ξ, it is easier to work with the orthogonal components $\left(w_A^+, w_B^+\right)$ of (w_A, w_B) in the value function iteration.

We evaluate the value function on an equi-spaced grid for the orthogonal components with a width of $\pm 4\sigma_{A,B}^+$ around their means, where $\sigma_{A,B}^+$ denote the long-run standard deviations of the orthogonal components. The grid is chosen to capture al-

[8] See for example Adda and Cooper (2003) for an overview of dynamic programming techniques.

[9] To save on notation we drop the regional index of a variable pair like (w_{At}, w_{Bt}) and just denote the pair by w_t.

most all movements of the income distribution F later on.[10] Given this grid, we can use Tauchen's (1986) algorithm to obtain the transition probabilities for the Markov-chain approximation of the income process in (17) .

We apply a multigrid algorithm (see Chow and Tsitsiklis, 1991) to speed up the calculation of the value function. This algorithm works iteratively. It first solves the dynamic programming problem for a coarse grid and then doubles the number of grid-points in each iteration until the grid is fine enough. In between iterations the solution for the coarser grid is used to generate the initial guess for the value-function iteration of the new grid. The initial grid has $16\times16\times32$ points (income A $\times$ income B $\times$ migration costs) and the final grid has 128×128 points for income and 256 points for migration costs.[11]

The solution of (6) yields the optimal migration policy and thus the microeconomic migration hazard rates Λ_i. With these hazard rates, we can obtain a series of aggregate migration rates for a simulated economy as described in detail in Section 4.2 for any realization of aggregate shocks $(\theta_t)_{t=1...T}$ and an initial distribution F_0.

This means that we need an initial distribution of income F_0 to solve the sequential problem. Following Caballero and Engel's (1999) suggestion, we use the ergodic distribution of income $\bar{F}$ that would be obtained in the absence of aggregate income shocks. This distribution is calculated by assuming that idiosyncratic shocks ω have the full variance of ξ. In the appendix, we show that the sequence of income distributions converges to a unique ergodic distribution $\bar{F}$ in the absence of aggregate shocks. This ergodic distribution $\bar{F}$ is a natural starting guess for F_0 as Caballero and Engel (1999) argue.

To simulate a series of migration rates which correspond to the aggregate migration hazards $\left(\bar{\Lambda}_{A,B}\right)_{t=1...T}$, we draw a series of aggregate shocks (to the orthogonal basis) $\left(\theta_A^+, \theta_B^+\right)_{t=1...T}$ from a normal distribution with variance $\phi \cdot \left(\sigma_{A,B}^+\right)^2$, $\phi \in [0,1]$. The

[10] The choice of $\pm4\sigma_{A,B}^+$ is motivated as follows. We later assume in the simulations that about 99% of the income shocks is due to the idiosyncratic component. Therefore, we can expect 99.9% of the mass of the income distribution to fall within $\pm3.29 \cdot \sqrt{0.99}\sigma_{A,B}^+ \cong \pm3.27\sigma_{A,B}^+$ around the mean of the distribution for any given year. Additionally, the mean income for each year moves within the band $\pm3.29 \cdot \sqrt{0.01}\sigma_{A,B}^+ \cong \pm0.33\sigma_{A,B}^+$ in again 99.9% of all years. Since the sum of both is $3.6\sigma_{A,B}^+$, a grid variation of $\pm4\sigma_{A,B}^+$ should not truncate the income distribution.

[11] To obtain the grid for migration costs, we first discretize the [0;1] interval into an equi-spaced grid. Then, we choose the grid points for the migration costs as the values of the inverse of the cumulative distribution function of the costs evaluated at the equi-spaced grid. This yields a cost grid whose grid points are equally likely to realize. By contrast to the income distribution, using such an "equally-likely grid" is possible for the cost distribution, because the cost distribution is strictly stationary. Unlike the income distribution, it does not move due to aggregate shocks. See Adda and Cooper (2003) or Tauchen (1986) for the analog case of a stationary Markov chain with normal innovations.

weight ϕ measures the relative importance of aggregate shocks, relative to idiosyncratic shocks, i.e. $\sigma_{\omega}^2 = (1-\phi)\,\sigma_{\xi}^2$ and $\sigma_{\theta}^2 = \phi\sigma_{\xi}^2$. Correspondingly, the orthogonal components of the idiosyncratic shocks have variance $(1-\phi)\cdot\left(\sigma_{A,B}^{+}\right)^2$.

5.2 Parameter choices

A number of parameters has to be determined to actually simulate our model numerically. Our parameter of most interest is migration costs. Our baseline specification of the model used for the simulations assumes migration costs to be Gamma-distributed, i.e. the cumulative distribution function of migration costs is

$$G(c) = \frac{1}{a^b \Gamma(b)} \int_0^c x^{b-1} \exp\left(\frac{-x}{a}\right) dx. \tag{18}$$

This distribution function has two parameters, a and b, which determine the mean ab and the coefficient of variation $b^{-\frac{1}{2}}$. Although the mean cost is ab, one should note that the average cost paid by a migrant can be smaller as she can wait and search for low migration costs. In our simulations, we try three parameter combinations (a, b) to see their influence on the dynamics of interregional migration. We try one parameter constellation with high, one with medium, and one with almost zero migration costs. We fix the coefficient of variation to 1 and choose mean costs to be US$ 180,000, US$ 45,000, and US$ 1, respectively. This allows us to assess the sensitivity of aggregate migration with respect to moving costs. In particular, we are interested to see whether the high migration-cost estimates reported in the literature are compatible with aggregate migration data in the light of our model.

As an alternative to this baseline specification of stochastic, Gamma-distributed migration costs, we also simulate the model with deterministic and constant costs of migration. This alternative specification implies that migration hazards $\Lambda_i(w_A, w_B)$ are either zero or one, now. Moreover, there is no longer an option value of searching for low migration costs that delays the migration decision in this simplified model. The only option value that the migrant takes into account is the value to wait for good income opportunities. When we estimate the model later on, we restrict our attention to this specification with deterministic costs, because in the more complex specification with stochastic migration costs, the two cost parameters a and b are only weakly identified separately.

The second important set of parameters describes the process for income and the income shocks ξ. We need to specify the autocorrelation parameter ρ and the mean μ of the income process as well as its covariance structure. The covariance structure is

composed of the total variance of income shocks σ^2_ξ, the correlation of income shocks between regions, ψ_θ (aggregate) and ψ_ω (idiosyncratic), and the fraction ϕ of the income shock that is due to aggregate factors, i.e. the correlation across individual agents.

We take the parameters for the income process mainly from the recent paper of Storesletten, Telmer, and Yaronx (2004). They estimate the dynamics of idiosyncratic labor market risk for the US based on the Panel Study of Income Dynamics. Thus the paper conveys information on both income variances and autocorrelation of log household income. Besides, the paper reports a mean household income of about US$ 45,000. To approximately match this figure, we choose the mean of the *log* income to be $\mu \cong 10.5$.[12]

Storesletten, Telmer, and Yaronx (2004) find an annual autocorrelation of incomes of roughly 0.95 and a standard deviation of idiosyncratic income shocks ranging from 0.09 to 0.14 for business cycle expansions and from 0.16 to 0.25 for business cycle contractions (see Storesletten, Telmer, and Yaronx 2004, Table 2). They report a frequency weighted average of 0.17 for those standard deviations in their preferred specification (Storesletten, Telmer, and Yaronx, 2004, pp. 711). Since we do not model different variances of idiosyncratic shocks to income along the business cycle, we use their preferred average value of 0.17 for the simulations.[13]

Combining both elements, the autocorrelation and the variance of idiosyncratic shocks to income, we calculate the long-run variance of income to be $\frac{\sigma^2_\omega}{1-\rho^2} = 0.30$. This number refers to persistent elements of income, which should be relevant to migration decisions. Of course, the fluctuation of income that is observed in real data does not only reflect these persistent shocks. Indeed, Storesletten, Telmer, and Yaronx (2004) find that transitory shocks to income add another variance term in the order of 0.065 to this long-run variance. This means that transitory shocks are responsible for about 18% of the total fluctuations of income. However, we expect these transitory shocks to be of minor relevance to migration choices, simply because they arrive at a too high frequency. Technically, we assume that the transitory shocks realize after migration decisions are taken and for this reason, we do not include any transitory components of income in the microeconomic model.

At the macroeconomic level, however, the inclusion of a transitory shock to income is of importance for two reasons if the model shall be compared to real-world data with respect to the correlation of incomes and migration rates.

[12] A log-normally distributed variable has mean $\exp\left(\mu + \frac{\sigma^2}{2}\right)$ where μ and σ^2 are the mean and variance of the logs.

[13] Other studies on the evolution of individual income report similar values, see the discussion in Storesletten et al. (2004).

Firstly, there will be some income fluctuations at the macro level that are transitory in a similar fashion as they exist at the micro level. This will influence the correlation of incomes and migration right away.

Secondly, and maybe more importantly, we have to take into account the fact that income measures migration incentives perfectly in our model while it obviously does not do so in the real world. For example, fluctuations in regional price levels, changes in the supply of public goods, or the fact that the empirical income concept is itself noisy, all together, weaken the relationship of income and migration at the aggregate level. This means that the model will produce unrealistically large correlations of income differentials and migration rates at the macroeconomic level if these aspects of measurement are ignored.

Both aspects, transitory income fluctuations and measurement problems, can be addressed by augmenting the model by a transitory error term of income.[14] For this reason, we introduce a such a term in the form of a pseudo normally distributed transitory shock to aggregate incomes. This transitory income component φ has no influence on the distribution of migration rates but only on the correlation of migration rates and incomes. We use the numbers reported by Storesletten, Telmer, and Yaronx (2004) for idiosyncratic shocks as an orientation. These numbers lead us to assume that the aggregate transitory shock φ has 18% of the long-run variance of the permanent aggregate income component, i.e. $\sigma_{\varphi}^2 = \frac{0.18\sigma_{\theta}^2}{1-\rho^2}$. This number seems to be a lower bound, however, because measurement errors should be an additional component of the transitory shock. For this reason, we estimate the magnitude of the transitory shock along with migration costs in the actual estimation of our model.

In order to describe the income process completely, two elements of the variance-covariance structure still have to be specified. We need to determine the magnitude of permanent aggregate fluctuations and the correlation of income shocks across regions. Unfortunately, Storesletten, Telmer, and Yaronx (2004) do not report numbers on aggregate income risk, so that we take this data from a different source. We estimate the variance of aggregate shocks to income from income per capita data for US states for the years 1969 - 2004 as reported in the REIS database of the BEA. This data is deflated using the US-wide consumer price index. Moreover, we remove fixed effects and a linear time trend from the income data. The residual variance of log income for US states over time is roughly 0.002.[15]

[14] This strategy picks up the idea of Erickson and Whited (2000) to rationalize empirically observed low investment-q sensitivities.

[15] Note that for the comparison of our model with real-world data we use a shorter time horizon to

To calculate the fraction, ϕ, of income risk due to aggregate fluctuations, we compare this estimated long-run aggregate variance with the long-run idiosyncratic variance of income that is implied by Storesletten, Telmer, and Yaronx's (2004) estimate of $\frac{\sigma_\omega^2}{1-\rho^2} = 0.30$. This number and and aggregate fluctuations of variance $\frac{\sigma_\theta^2}{1-\rho^2} = 0.002$ implies an overall long-run variance of income of 0.302. In turn, aggregate income risk accounts only for a fraction of approximately $\frac{0.002}{0.3+0.002} \cong 0.006$ of total income risk. For the simulations, we use this number to specify ϕ. However, our rough calculation of the fraction of aggregate shocks can only be an approximation. Therefore, we actually estimate this fraction later on.

Finally, we need to specify the correlations of shocks to income across regions, ψ_ω and ψ_θ. These correlations refer to *potential* incomes and are therefore inherently unobservable. We assume that transitory, aggregate, and individual correlation coefficients are equal, i.e. $\psi_\omega = \psi_\theta = \psi_\varphi$, so that we only need to specify one common parameter. In our simulation exercise, we measure ψ as the correlation coefficient of state-average income per capita and the US-average per capita income (both in logs, CPI deflated and taking fixed effects and a linear time trend into account). From the REIS database, we infer a partial correlation coefficient of $\hat{\psi} = 0.578$. Again, this number can only be a first approximation. For the estimation, we abstain from fixing the parameter ψ but estimate it along with migration costs.

As we work with annual data, we choose the discount factor $\beta = 0.95$. Table 1 summarizes our parameter choices for the three specifications that we simulate.

5.3 Simulation results

We simulate our model for 51 pairs of regions and 26 years, but we drop the first 10 years for each region to minimize the influence of our initial choice of F_0. This generates a simulated dataset for migration that has the same size as the Internal Revenue Service (IRS) area-to-area migration flow dataset, which is our empirical benchmark. This data base contains area-to-area migration flow data for US states for the years 1989-2004. Income data is taken from the REIS database, CPI deflated, and in logs. A detailed data description for both IRS and REIS data can found in the data appendix. In order to minimize simulation uncertainty, we replicate each simulation 10 times and report the averages over the simulations.

Of course the actual migrant faces a more complex decision problem than the one

calculate summary statistics. This is done to match the length of the IRS data. This implies that the within sample variance of aggregate income presented in these summary statistics is smaller than the estimate of the long-run variance presented here.

Table 1: Parameter choices for the simulation analysis

	Storesletten et al. (2004)	REIS data	Simulation
Fraction of aggregate shocks ϕ	–	0.006	0.006
Correlation of shocks across regions ψ	–	0.578	0.578
Long-run variance of incomes $\frac{\sigma_\omega^2+\sigma_\theta^2}{1-\rho^2}$	0.30	0.002	0.302
Variance of transitory aggregate income shock	–	–	3.61×10^{-4}
Autocorrelation of income ρ	0.95	–	0.95
Discount factor β	–	–	0.95

simulated in our model of two regions. Including D.C. as a destination region, an agent has to decide between 50 possible alternatives states where she can move to. To make this comparable to our model, the 50 alternatives in the data have to be aggregated to a single complementary region.[16] The population-weighted average income over all alternative 50 states is used as the average income of the alternative region.

In order to characterize the results of the simulation exercise, we have to calculate a number of moments from the simulated dataset and compare these moments to the moments that we observe in the actual IRS and REIS data. This comparison tells us how well our model is capable of replicating characteristic features of the actual migration and income data at an aggregate level. In particular, the comparison tells us which of the three considered levels of migration costs is best compatible with the observed data. Such way of inference is frequently applied in the literature on real business cycles, see Backus, Kehoe, and Kydland (1992) or Baxter and Crucini (1993) and many others.

The lines along which this literature has typically described aggregate fluctuations guide our choice of characterizing moments: variances, covariances, autocorrelations, and

[16] Generating artificial bi-regional data means that we assume technically that the best income opportunity over all alternative regions follows the log-normal distribution assumed in our model.

means. We compare average migration rates, the standard deviation of migration rates, their autocorrelation, and the cross-correlation of migration rates. Besides, we look at the implications of the different migration cost regimes on the level and fluctuations of average incomes. To measure the cyclical behavior of migration, we calculate the mean of in- and outmigration rates and correlate this with the average income in both regions. In the simulated data, the average income in region i is calculated $\bar{w}_{it} := \ln\left(\int \exp w_i \hat{f}_{it}\left(w_i, w_{-i}\right) dw_i dw_{-i}\right)$, i.e. the income obtained after migration decisions have been taken. Additionally, we run a typical reduced form migration regression that relates migration rates to the average incomes in source and destination region.

We run two sets of simulation exercises. One set features stochastic migration costs that are Gamma-distributed. The other set of simulation exercises assumes migration costs to be deterministic and constant over time. While the first formulation may be regarded as being more realistic, since migration costs are modelled more flexible at the household level, the corresponding simulation results are harder to interpret at the same time. Stochastic migration costs and the household's option to search for low costs of migration drive a wedge between the expected level of migration costs and the costs that are actually incurred by the migrating households. In other words, the first set of simulations involves two different measures of migration costs. This makes our analysis somewhat difficult to relate to findings of previous studies, which work with a deterministic formulation of migration costs. This is the reason why we also report a set of simulation exercises in which migration costs are fixed to a deterministic value.

Table 2 reports the results of our first set of simulation exercises, in which migration costs are stochastic and follow a Gamma-distribution. The first experiment uses cost parameters close to what has been reported in the literature. We fix the coefficient of variation in all three experiments to one. To match an average migration cost of US$ 180,000 as reported in Davies, Greenwood, and Li (2001), we set $a = 180,000$. The results of this experiment are displayed in column (1) of Table 2.

Compared to the actual data, the annual migration rates are too low. While we observe an annual average migration rate of 3.9%, the model predicts a migration rate of only 2.9%. With US$ 180,000, expected migration costs are too high. Also migration rates fluctuate less in the simulated data than in the actual data. Simulated migration rates are too much procyclical and the cross-correlation of incomes is 0.611, while the correlation of income shocks ψ was set to be 0.578. With migration costs being stochastic, the actually incurred migration costs are with about US$ 23,000 much lower than the expected migration costs. This difference stems from the substantial variation of migration costs which was imposed by our ad hoc choice of $b = 1$. In turn, potential

Table 2: Simulation results: stochastic, Gamma-distributed migration costs

	Data	high costs (1)	medium costs (2)	zero costs (3)
Average annual migration rate	0.039	0.029	0.041	0.102
Standard deviation of annual migration rates	0.004	0.002	0.004	0.009
Autocorrelation of migration rates[1]	0.807	0.839	0.739	0.279
Cross correlation of migration rates[1]	0.047	-0.972	-0.979	-0.991
Mean of log average income	10.710	10.825	10.836	10.833
Standard deviation of log average income	0.030	0.027	0.026	0.025
Cross-correlation of log average income [1]	0.578	0.611	0.652	0.737
Variance of household income	0.299	0.295	0.287	0.261
Correlation of $\left(\bar{\Lambda}_i + \bar{\Lambda}_{-i}\right)$ and $(\bar{w}_i + \bar{w}_{-i})$ (procyclicality)[1]	0.215	0.540	0.467	0.016
Sensitivity of immigration into region i [2]				
w.r.t. average income in region i	0.061	0.068	0.103	0.111
w.r.t. average income in region $-i$	-0.063	-0.063	-0.097	-0.110
Average incurred migration costs		23,043	10,675	0.99

[1] Partial correlation controlling for a linear time trend and fixed effects.
[2] Coefficients of a reduced form regression of migration rates on incomes in both regions.
All three specifications assume a coefficent of variation of 1 for migration costs, i.e. $b = 1$. The high cost specification assumes expected migration costs to be US$ 180,000, i.e. $a = 180,000$. The medium cost specification assumes $a = 45,000$ and the zero cost specification sets $a = 1$ for numerical feasiblity. We simulate data on 51 region-pairs and a 26 year history of migration and income data. The first 10 years of simulated data are dropped in order to minimize the influence of initial values. Each simulation is repeated 10 times.

migrants wait for low realizations of migration costs which are drawn every period anew.

In summary, the high cost specification implies too little migration and too little fluctuation of the migration rates, while income fluctuation is realistic. Therefore, we try a specification with lower migration costs. We set $a = 45,000$, so that expected migration costs are divided by four and now equal an average annual income of US$ 45,000. With these lower migration costs, migration rates increase substantially and are with 4.1% very close to the observed average migration rate. Also the standard deviation of migration rates is very close to the one we observe in the data. Migration also becomes less procyclical, but aggregate migration responds overly strong to aggregate income. A further result of lower migration costs is an increase in average income by 1.1% compared to the high cost specification. With lower migration costs, the agents are more often in the region where their income is larger. However, households are in their preferred region most of the time already in the high cost specification, so that potential benefits of further migration are small. This reflects that the actually *incurred* migration costs are relatively small already in the specification with high costs. In the specification with medium expected migration costs, the incurred costs are even smaller and only amount to US$ 10,675.

While the first scenario displayed an extreme bound of high migration costs, the third scenario of almost no migration costs provides a lower bound. It clearly shows how influential it is to keep track of the evolution of migration incentives. In a model in which migration incentives are drawn randomly, we should observe migration rates of 50% in the absence of migration costs. By contrast, our model predicts a substantially lower migration rate of 10.2% when migration costs are absent. This difference stems from the fact that in our model migration incentives are not drawn purely randomly. Instead, they depend on previous migration decisions and income shocks.

The differences between the three cost specifications become even more pronounced when we assume migration costs to be deterministic, see Table 3. Qualitatively the results do not change when apply this simplification. Quantitatively, the differences between the three specifications become more pronounced though. Migration rates are far too low in the high cost specification and they fluctuate way too little. The efficiency gain due to better allocation of households to regions measured by the average income is substantially larger when migration costs are reduced from US$ 180,000 to US$ 45,000. Here, the average income increases by 3.7%.

Overall, our various simulation exercises do not yet allow a decisive assessment of which level of migration costs fits the data best. The average migration rates and their fluctuations are best captured by the medium cost formulation. However, the overall

Table 3: Simulation results: deterministic migration costs

	Data	high costs (1)	medium costs (2)	zero costs (3)
Average annual migration rate	0.039	0.010	0.024	0.102
Standard deviation of annual migration rates	0.004	0.001	0.003	0.009
Autocorrelation of migration rates[1]	0.807	0.687	0.633	0.279
Cross correlation of migration rates[1]	0.047	-0.785	-0.941	-0.991
Mean of log average income	10.710	10.784	10.821	10.833
Standard deviation of log average income	0.030	0.027	0.026	0.025
Cross-correlation of log average income [1]	0.578	0.560	0.616	0.737
Variance of household income	0.299	0.310	0.277	0.261
Correlation of $\left(\bar{\Lambda}_i + \bar{\Lambda}_{-i}\right)$ and $(\bar{w}_i + \bar{w}_{-i})$ (procyclicality)[1]	0.215	0.550	0.457	0.016
Sensitivity of immigration into region i [2]				
w.r.t. average income in region i	0.061	0.033	0.078	0.111
w.r.t. average income in region $-i$	-0.063	-0.024	-0.068	-0.110
Average incurred migration costs		180,000	45,000	1

[1] Partial correlation controlling for a linear time trend and fixed effects.
[2] Coefficients of a reduced form regression of migration rates on incomes in both regions.
All three specifications assume deterministic migration costs. The high cost specification assumes migration costs to be US$ 180,000. The medium cost specification assumes migration costs of US$ $45,000$ and the zero cost specification assumes one US$ of migration costs for numerical feasiblity. We simulate data on 51 region-pairs and a 26 year history of migration and income data. The first 10 years of simulated data are dropped in order to minimize the influence of initial values. Each simulation is repeated 10 times.

match of the simulated data with the observed data is not perfect.

6 Estimation

We rely on an indirect inference procedure in order to find the parameters of our model that allow us to match closest the observed patterns of migration that are in the data. In particular, we apply a method of simulated moments (MSM) as has been proposed by Gourieroux, Monfort, and Renault (1993) to obtain estimates of structural parameters when the likelihood function of the structural model becomes intractable, as in our setting.

6.1 Methodology

Indirect inference is the natural extension of the simulation exercise presented in the previous section. The idea behind this methodology is to choose a set of moments that captures the characteristics of the data, and then to calibrate and simulate the structural economic model such that the moments are best replicated by the simulation.

Accordingly, we first decide on an informative set of moments ϱ. We select the standard deviation of migration rates, the standard deviation of average incomes, and the correlation of average incomes across regions as the first three moments to be matched. To this set of moments we add the estimated parameters from a reduced form regression of migration rates on the incomes of the destination and the source region. To make the regression scale invariant with respect to incomes, we use log-deviations from average incomes as the income variables, i.e. we estimate

$$m_{it} = \alpha_0 + \alpha_1 \left(w_{it} - \bar{w}_i \right) + \alpha_2 \left(w_{-it} - \bar{w}_{-i} \right) + u_{it}. \tag{19}$$

The parameters α_1 and α_2 reflect income sensitivities of migration. For the simulated data, these sensitivies were reported lower parts of Tables 2 and 3. The intercept α_0 captures the average of migration rates.

We simulate our model for a given vector of model parameters β and calculate the distance between the moments obtained from this simulation $\hat{\varrho}\left(\beta\right)$ and the sample moments ϱ_S. We use the covariance matrix of ϱ_S obtained by 10000 bootstrap replications as a weigthing matrix so that our distance and goodness-of-fit measure is

$$L = \left(\varrho_S - \hat{\varrho}\left(\beta\right) \right)' \operatorname{cov}\left(\varrho_S\right)^{-1} \left(\varrho_S - \hat{\varrho}\left(\beta\right) \right). \tag{20}$$

Naturally, we cannot estimate all parameters of the model, since this would be numerically infeasible. We restrict ourselves to the estimation of migration costs and the

correlation of shocks to *potential* incomes both across individuals (ϕ) and across regions (ψ). We opt for the estimation of ϕ and ψ, because these parameters cannot be inferred from *realized* income data alone as we argued previously. Since migration smoothes income, the counterparts to ϕ and ψ in terms of a covariance structure in realized incomes are substantially influenced by the size of migration costs. At the same time, we expect ϕ and ψ to have a significant influence on the behavior of aggregate migration itself. However, this argument does not hold true for the correlation of the transitory aggregate income shock φ. This transitory shock is irrelevant to the migration decision itself, and hence its effect cannot be smoothed by migration. For this reason, we fix the correlation of the transitory aggregate income shock ψ_φ at the value of the observed correlation of incomes in the REIS data.

While we have tried both a stochastic as well as a deterministic specification of migration costs in the simulations, we restrict ourselves to the estimation of a specification with deterministic costs for two reasons. First, the formulation with deterministic cost is much easier to interpret and to compare to other studies, as migration costs are captured by a single number to be estimated. Second, the two parameters of the Gamma-distribution in a formulation with stochastic costs are only weakly identified separately. Indeed, we find that the formulation with stochastic costs increases the risk of running into local minima of the distance measure L.

We estimate σ_φ^2 along with the other model parameters, so that our set of estimated parameters finally is $\beta = \left(\text{migration costs}, \psi, \phi, \sigma_\varphi^2\right)$.[17]

6.2 Estimation results

Table 4 displays the point estimates of the matched moments calculated from the IRS and REIS data and the corresponding moments obtained from the simulation of our model with the estimated parameters. Overall our model is able to replicate the observed moments closely. In fact, the overidentification test reported at the bottom of Table 5 does not reject our model.

Table 5 presents the estimates of the model parameters and the $\chi^2(2)$-distributed overdentification test. The estimated migration costs are US$ 19,707. This number lies below the costs considered in the previous section for the medium specification with deterministic costs. It falls between the average incurred costs of the medium and the high cost-specification with stochastic costs. In any case, it is a smaller number than the estimates reported in other contributions such as Davies, Greenwood, and Li (2001)

[17] To save on computation time, we use a smaller grid than in the simulation exercises. We choose a grid of $64 \times 64 \times 128$ points to approximate the state space.

Table 4: Simulated moments estimation: moments estimates

Moment	Actual Moments	Simulated Moments
Migration rates		
standard deviation	0.0036	0.0036
Income		
standard deviation	0.0299	0.0297
correlation across regions	0.5776	0.5646
Reduced form regression		
intercept (average migration rate)	0.0393	0.0392
sensitivity to destination income	0.0609	0.0617
sensitivity to source region income	-0.0627	-0.0604

The column 'Actual Moments' refers to the moments estimated from the combined REIS/IRS data set, with data on 50 US states and D.C. over the period 1989-2004. The column 'Simulated Moments' refers to the moments estimated from the simulation of the model using the parameters given in Table 5. Both actual and simulated data are within-transformed and linearly de-trended. The simulations generate a panel of 51 region-pairs and a 26-year history of migration and income data. The first 10 years of simulated data are dropped in order to minimize the influence of initial values. Each simulation is repeated 5 times and data moments are compared to the average over the 5 replications of the simulation.

or Kennan and Walker (2006).

The estimated value of the correlation of income shocks across regions is 0.2507. This is substantially smaller than we specified for the simulations, having set the correlation of shocks equal to the observed correlation of realized incomes (0.5776, see Table 1). Migration ties together more closely the average incomes of both regions than they were tied together without migration. The realized incomes co-move more strongly than the shocks to the income process. This drives a wedge between the correlation of income shocks and the correlation of average realized incomes, the latter being always larger than the former.

The estimated fraction of income shocks that is aggregate amounts to 0.0043. This size of the aggregate shock corresponds closely to the value we previously assumed. There is a significant transitory income component in the aggregate income fluctuations, which has an estimated standard deviation of 0.0265. This means that transitory fluctuations in aggregate income add a variance term that has about 54% of the long-run variance of potential incomes $\left(\frac{0.0265^2}{0.0043\times 0.302} = 0.5408\right)$. However, migration smoothes realized income, so that transitory shocks make up 78% of the aggregate variance in realized income.

As outlined before, the transitory income component of our model relates to two sources: to truely transitory fluctuations in incomes and to the fact that migration is not perfectly driven by income incentives alone. The latter aspect refers to other aggregate factors that are important to migration flows and were subsumed as income in our model.

6.3 Comparison of cost estimates

To provide further evidence on the influence of the dynamics of the incentive distribution on the estimation of migration costs, we apply a static random utility model to data generated from a simulation of our dynamic model. The generated data set comprises 51 pairs of regions and 16 years of data. The parameters of the model are fixed to the values as estimated in the previous section.

The aim of this exercise is to facilitate a direct comparison of static and dynamic approaches to the estimation of migration costs. In particular, we apply a conditional-logit approach similar to Davies, Greenwood, and Li (2001) to describe the migration decision. Simplifying Davies, Greenwood, and Li's model and adapting it to our bi-regional framework, the likelihood of the conditional logit model becomes

Table 5: Simulated moments estimation: structural parameter estimates

Parameter	Estimated value
Migration costs	19707 (655.49)
Correlation of income shocks across regions Ψ	0.2507 (0.0103)
Fraction of income shock due to aggregate fluctuations ϕ	0.0043 (0.0004)
Standard deviation of transitory income shock ψ_{φ}	0.0265 (0.0010)
$\chi^2(2)$ p-value	0.9593 0.619

Standard errors in parenthesis. Estimation is carried out using the simulated moments estimator in Gourieroux, Monfort, and Renault (1993), which chooses structural model parameters by matching the moments from a simulated panel of regions matching the moments as displayed in Table 4. The simulations generate a panel of 51 region-pairs and a 26-year history of migration and income data. The first 10 years of simulated data are dropped in order to minimize the influence of initial values. Each simulation is repeated 5 times and data moments are compared to the average over the 5 replications of the simulation.

Table 6: Simulation results: comparison to cost estimate based on a static random utility model

Migration cost (from estimation)	19,707
Average annual income	53,906
Migration cost estimate based on a static random utility model	102,930

$$\ln L = \sum_t \sum_{i=1,2} \left[\begin{array}{c} \bar{\Lambda}_{it} P_{it} \ln \left(\frac{1}{1+\exp\{c+\gamma(\bar{w}_{it}-\bar{w}_{-it})\}} \right) \\ + \left(1-\bar{\Lambda}_{it}\right) P_{it} \ln \left(\frac{1}{1+\exp\{-c-\gamma(\bar{w}_{it}-\bar{w}_{-it})\}} \right) \end{array} \right]. \tag{21}$$

While Davies, Greenwood, and Li (2001) include a set of other variables to describe the utility gained from location choice, our simulated model just allows for log income as an explanatory variable. This means that the form of the likelihood function in (21) assumes that utility is composed of an income component (with sensitivity $\gamma > 0$) and a disutility from migration $c < 0$. The estimated money measure of this disutility is $\exp\left(\bar{w} - \frac{\hat{c}}{\hat{\gamma}}\right)$, see Davies, Greenwood, and Li (2001). Since our model is composed of two regions only, we cannot estimate γ and c from a cross-section as Davies, Greenwood, and Li (2001) do, but have to pool the simulated data instead.

We want to abstract from the additional problem that income data measures migration incentives imperfectly. Otherwise, this would drive up estimated migration costs and bias the comparison against the static approach. For this reason, we simulate the dynamic model having set all parameters to their estimated values except for the variance of the measurement error, which is set to zero instead.

The imputed migration costs taken from our estimation are US$ 19,707. By contrast, the conditional logit estimation suggests a cost of US$ 102,930, a number that is substantially higher (see Table 6). In terms of annual incomes this corresponds to 0.4 and 2 average annual incomes, respectively. This comparative exercise shows that the estimation of the structural parameters is likely to be subject to a bias if the unobserved dynamics of the distribution of incentives is not taken into account.

7 Conclusion

We have provided a tractable model of aggregate migration with a sound microeconomic foundation. The paper is a contribution to the recently evolving literature on structural models of migration. We explicitly deal with the problem of the unobservability of potential gains from migration and their dynamic character. The dynamic character of migration incentives has two aspects. First, the individual gains from migration evolve stochastically over time, but will typically display a high degree of persistency. Second, at an aggregate level, the distribution of migration incentives is a result of past migration decisions themselves.

Starting off from the microeconomic decision problem allows us to keep track of the dynamics of the incentive distribution. This distributional dynamics may be refered to as a dynamic self-selection problem. Neglecting this self-selection problem may result in biased estimates of structural parameters, such as migration costs. In our application to US interstate migration, we find the estimated migration costs to be substantially lower than those reported in previous studies. The estimated migration costs amount to about US$ 20,000, which corresponds to less than one-half of an average annual income.

Our analysis calls once more for a careful treatment of the self-selection problem when economic incentives are not fully observable. What makes this issue particularly relevant for the analysis of migration is that the unobservable incentives are highly autocorrelated though not perfectly persistent. Rather than being drawn every period anew, migration incentives have a long memory. One example of this long memory of migration incentives is the persistency that income displays.

This may be of importance not only to macro-studies of migration. Also at a micro level, income potentials are typically unobservable and have to be proxied. But such approximation regularly neglects self-selection. If households live in their prefered place of residence as a result of their location choice, and if all observable things are equal, then it must be the unobserved component of their preferences that is in favor of the place where they actually are. Besides unobservable parts of income, this unobservable component of preferences can also comprise different valuations of different amenities and social networks. Also these factors can be expected to exhibit persistency.

We integrated the persistency of unobserved migration incentives in a structural dynamic microeconomic model of the migration decision. This consequently allowed us to simulate the joint behavior of the observed migration rates, of the unobserved migration incentives, and of their observable proxy, incomes.

Future research calls for a more complex microeconomic model that integrates more

information into the macroeconomic analysis, for example labor market conditions and amenities. Additionally, it would be desirable to extend our bi-regional approach to the case of multiple regions, as in Davies, Greenwood and Li (2001), and Kennan and Walker (2006). Ultimately, migration should be analyzed in a more complex general equilibrium framework. However, all this goes beyond what is currently numerically feasible, in particular if the model is meant to be estimated.

Both our treatment of the self-selection problem and the inference of microeconomic structural parameters from macroeconomic data is an attempt to overcome the dichotomy of macro and micro studies that has characterized the migration literature (see Greenwood, 1997). Beyond the application to migration decisions, our treatment of the dynamic self-selection problem may also be applicable to other important discrete choices in an economy, for example labor-market participation.

8 Appendix

8.1 Existence and uniqueness of the value function

We begin with proving existence and uniqueness of the value function. Notation is as in the main text throughout this appendix, unless stated otherwise.

To ease the exposition, we assume that the income process is only approximately log-normal. In particular, we assume that income has a finite support.

Definition 1 *Let* $\mathbb{W} = \left[\underline{W}, \overline{W}\right]$ *be the support of* w.

Definition 2 *Define a mapping* T *according to the migration problem of a household, that is*

$$T(u)(\cdot) = \max_{j=A,B} \left\{ \exp(w_{jt}) - \mathbb{I}_{\{i \neq j\}} c_t + \beta \mathbb{E}_t u(j, c_{t+1}, w_{At+1}, w_{Bt+1}) \right\}. \tag{22}$$

The mapping T *is defined on the set of all real-valued, bounded functions* $\mathbb{B}$ *that are continuous with respect to* $w_{A,B}$ *and* c *and have domain* $D = \{A, B\} \times \mathbb{R}_+ \times \mathbb{W}^2$.

Lemma 3 *The mapping* T *preserves boundedness.*

Proof. *To show that* T *preserves boundedness one has to show that for any bounded function* u *also* Tu *is bounded. Consider* u *to be bounded from above by* $\bar{u}$ *and bounded from below by* $\underline{u}$*. Then,* Tu *is bounded, because*

$$Tu = \max_{j=A,B} \left\{ \exp(w_{jt}) - \mathbb{I}_{\{i \neq j\}} c_t + \beta \mathbb{E}_t u(j, c_{t+1}, w_{At+1}, w_{Bt+1}) \right\} \leq \exp\left(\bar{W}\right) + \beta \bar{u} < \infty, \tag{23}$$

and

$$Tu = \max_{j=A,B} \left\{ \exp\left(w_{jt}\right) - \mathbb{I}_{\{i \neq j\}} c_t + \beta \mathbb{E}_t u\left(j, c_{t+1}, w_{At+1}, w_{Bt+1}\right) \right\} \tag{24}$$

$$\geq \max_{j=A,B} \left\{ \exp\left(w_{jt}\right) - \mathbb{I}_{\{i \neq j\}} c_t + \beta \underline{u} \right\} \geq \exp\left(\underline{W}\right) + \beta \underline{u} > -\infty. \tag{25}$$

■

Lemma 4 *The mapping T preserves continuity.*

Proof. *Since Tu is the maximum of two continuous functions it is itself continuous.* ■

Lemma 5 *The mapping T satisfies Blackwell's conditions.*

Proof. *First we need to show that for any $u_1(\cdot) < u_2(\cdot)$ the mapping T preserves the inequality. Since both the expectations operator and the* max *operator preserve the inequality, also T does. Secondly, we need to show that $T(u + a) \leq Tu + \gamma a$ for any constant a and some $\gamma < 1$. Straightforward algebra shows that*

$$T(u + a) = Tu + \beta a. \tag{26}$$

Since $\beta < 1$ by assumption, T satisfies Blackwell's conditions. ■

Proposition 6 *The mapping T has a unique fixed point on $\mathbb{B}$, and hence the Bellman-equation has a unique solution.*

Proof. *Follows straightforwardly from the last three Lemmas.* ■

8.2 Invariant distribution

We prove that migration and idiosyncratic shocks to income describe an ergodic Markov-process if there are no aggregate shocks. Therefore, there is an invariant distribution the sequence of income distributions converges to.

For simplicity, we present the proof for an arbitrary discrete approximation of the model with a continuous state-space for income.

Lemma 7 *Assume any (large and fine enough but otherwise arbitrary) discretization of the state space with n points for the potential income in the regions, each. Then, we can capture the transition from f_t to f_{t+1}, which are the unconditional densities of the distribution of households over both regions and potential incomes, in a matrix*

$B = \begin{pmatrix} (I - D_A)\,\Pi & D_B\Pi \\ D_A\Pi & (I - D_B)\,\Pi \end{pmatrix} \in R^{2n^2 \times 2n^2}$.[18] *In this matrix,* Π *denotes the transition matrix that approximates the AR(1)-process for income by a Markov-chain, see Adda and Cooper (2003, pp. 56) for details. Matrix* D_i *is the* $n^2 \times n^2$ *diagonal matrix with the migration hazard rates for each of the* n^2 *income pairs of the income grid.*

Proof. *First, we take a discrete state-space of* n *possible wages for each region,* $w_{A1}...w_{Bn}$ *and* $w_{B1}...w_{Bn}$*. Second, we denote the vector of probabilities that describes the distribution of potential incomes and household locations in the following form*

$$f = \Big(f\left(A, w_{A1}, w_{B1}\right) ... f\left(A, w_{An}, w_{B1}\right) ... f\left(A, w_{An}, w_{Bn}\right) f\left(B, w_{A1}, w_{B1}\right) ... f\left(B, w_{An}, w_{Bn}\right) \Big)'. \tag{27}$$

Analogously, we define the distribution after migration but before idiosyncratic shocks, $\hat{f}$*. Taking our law of motion from* (16)*, we obtain as a discretized analog*

$$f_{t+1} = \left(I_2 \otimes \Pi\right) \hat{f}_t. \tag{28}$$

Here $\otimes$ *denotes the Kronecker product. Now, define* d_i *as the fraction of households that migrate and are in the* $i-th$ *income and location triple given our vectorization of the income grid. This means that* $d_i = \Lambda_j\left(w_{Ak}, w_{Bl}\right)$*,* $i = 1...2n^2$*, where* (j, w_{Ak}, w_{Bl}) *being the i-th element in the vectorized grid. Moreover, define* $D = diag\left(d\right)$ *as the diagonal matrix with migration rates on the diagonal and* D_A *and* D_B *as the diagonal matrices with only the first* n^2 *and the last* n^2 *elements of* d*, respectively. Then, we can describe the transition from* f_t *to* $\hat{f}_t$ *by*

$$\hat{f}_t = \begin{pmatrix} I - D_A & D_B \\ D_A & I - D_B \end{pmatrix} f_t \tag{29}$$

Combining the last two equations, we obtain

$$f_{t+1} = \begin{pmatrix} (I - D_A)\,\Pi & D_B\Pi \\ D_A\Pi & (I - D_B)\,\Pi \end{pmatrix} f_t. \tag{30}$$

■

Lemma 8 *For any distribution of idiosyncratic shocks with support equal to* $\mathbb{W}^2$*, matrix* Π *has only strictly positive entries.*

[18] Since we work with a discretization, correctly speaking f is not the density, but the vactor of probabities of drawing a location-income possibility vector from a given element of the grid.

Proof. *If the idiosyncratic shocks have support equal to* $\mathbb{W}^2$*, then every pair of potential incomes can be reached from every other pair as a result of the shock, because we assume the shocks to incoem being approximately log-normal. Thus, all entries of* Π *are strictly positive.* ■

Lemma 9 *For any distribution of costs with support equal to* $\mathbb{R}_+$*, the inequalities* $0 \leq d_i < 1$ *hold for all diagonal elements* d_i *of* D*. If the grid is fine enough also* $d_i > 0$ *holds at least for one* i*.*

Proof. *If there is no upper bound to migration costs, the migration probability is strictly smaller than 1, since* V *is bounded. This means* $0 \leq d_i < 1$*. Let* $C_{\max} = \max_{(w_A,w_B)\in\mathbb{W}^2} |\bar{c}(w_A, w_B)|$ *be the largest possible gain from migration. If the grid for costs is fine enough, there will always be a grid-point of migration costs that is smaller than this maximal gain* $C_{\max}$*. This is because migration costs can be arbitrarily close to zero. Hence, there is some* i *such that* $d_i > 0$ *holds if the grid is fine enough.* ■

Lemma 10 *For any distribution of costs with support equal to* $\mathbb{R}_+$*,* B^2 *has only positive entries.*

Proof. *We obtain for* B^2

$$B^2 = BB = \begin{pmatrix} ((I-D_A)\Pi)^2 + D_B\Pi D_A\Pi & (I-D_A)\Pi D_B\Pi + D_B\Pi(I-D_B)\Pi \\ (I-D_B)\Pi D_A\Pi + D_A\Pi(I-D_A)\Pi & ((I-D_B)\Pi)^2 + D_A\Pi D_B\Pi \end{pmatrix}. \tag{31}$$

Each entry of this matrix is weakly positive, because $(I-D_i)$*,* D_i*, and* Π *are positive. Hence, we only need to argue that in each sum at least one part is always strictly positive. For the elements on the diagonal, this follows directly from* $(I-D_i)\Pi > 0$*. For the off-diagonal elements, there may be some rows of zeros in* $D_i\Pi$*. However, at least one row of* $D_i\Pi$ *will be non-zero, because there is some non-zero* d_i *and* $(I-D_i)\Pi > 0$ *because of the Lemma above. Consequently, all elements of* $(I-D_i)\Pi D_j\Pi$ *are strictly positive.*

Proposition 11 *Under the assumptions of the above Lemmas, migration and idiosyncratic shocks define an ergodic process with stationary distribution* $F_0 = \lim_{n\to\infty} B^n e_i$*.*

Proof. *The above Lemma directly implies the ergodicity of the Markov chain.* ■

■

8.3 Data

Data on migration between US states are provided by the US Internal Revenue Service (IRS). The IRS uses individual income tax returns to calculate internal migration flows

between US states. In particular, the IRS compiles migration data by matching the Social Security number of the primary taxpayer from one year to the next. The IRS identifies households with an address change since the previous year, and then totals migration to and from each state in the US to every other state. Given these bilateral migration data, we compute aggregate gross immigration for the 50 US states and the District of Columbia as the sum of all immigrations from other US states to a particular state. Migration rates are calculated by expressing gross immigration as proportions of the number of non-migrants reported in the IRS dataset. The IRS state-to-state migration-flow data is available for the years 1989-2004.

Income per capita data are taken from the Regional Economic Information System (REIS) compiled by the Bureau of Economic Analysis. The REIS data is available online at www.bea.gov/bea/regional/reis/. The income-per-capita figure for the alternative region is computed as the population-weighted mean of all per-capita incomes outside a specific state.

We remove a linear time trend from all data and express all variables as deviations from their unit-specific means (rescaled by their overall mean), i.e. we apply a within-transformation. Table 7 reports descriptive statistics for the original as well as for the transformed data.

In order to examine the time-series properties of the data employed, we perform a unit-root analysis for migration rates and income data. In a sample of this size ($T = 16, N = 51$) either a Breitung and Meyer (1994) or a Levin, Lin, and Chu (2003) unit-root test are most appropriate. For the Breitung and Meyer (1994) test, we determined the optimal augmentation lag length by sequential $t-$testing. Taking into account three augmentation lags and time-specific effects we can reject the null hypothesis of a unit root at the 5% level of significance. Similarily, the Levin, Lin, and Chu (2003) test rejects the null hypothesis of a unit root taking a linear time trend into account.

References

[1] Adda, J. and R. Cooper(2003): "Dynamic Economics: Quantitative Methods and Applications", MIT Press, Cambridge.

[2] Baxter, M. and M. Crucini (1993): "Explaining saving/investment correlations", American Economic Review, 83, 416-436.

[3] Backus, D., P. Kehoe, and F. Kydland (1994): "Dynamics of the trade balance and the terms of trade: the J-curve?", American Economic Review, 84, 84-103.

[4] Blanchard, O. and L. Katz (1992): "Regional evolutions", Brookings Papers on Economic Activity, 1, 1-75.

Table 7: Descriptive statistics

	Mean	Std. Dev.	Min	Max
migration rate	.0393	.0178	.0144	.1146
migration rate filtered	.0393	.0036	.0234	.0629
income per capita	10.71	.1644	10.39	11.27
income per capita filtered	10.71	.0312	10.61	10.82
complementary income per capita	10.76	.0624	10.67	10.86
complementary income per capita filtered	10.76	.0239	10.73	10.80

[5] Breitung, Jörg and Wolfgang Meyer (1994): "Testing for unit roots in panel data: are wages on different bargaining levels cointegrated?", Applied Economics, 26, 353-361.r

[6] Burda, M. (1993): "The Determinants of East-West German Migration: Some First Results", European Economic Review, 37, 452-462.

[7] Burda, M., M. Müller, W. Härdle, and A. Werwatz (1998): "Semiparametric Analysis of German East-West Migration Intentions: Facts and Theory", Journal of Applied Econometrics, 13, 525-541.

[8] Borjas, G. J. (1987): "Self-selection and the earnings of immigrants", American Economic Review, 77, 531-553.

[9] Borjas, G. J., S. Bronars, and S. Trejo (1992): "Self-selection and internal migration in the United States", Journal of Urban Economics, 32, 159-185.

[10] Caballero, R. and E. M. R. A. Engel (1999): "Explaining Investment Dynamics in U.S. Manufacturing: A Generalized (S,s) Approach", Econometrica, 67, 783-826.

[11] Chow, C. S. and J. N. Tsitsiklis (1991): "An optimal multigrid algorithm for continuous state discrete time stochastic control", IEEE Transactions on Automatic

Control, 36, 898–914.

[12] Cushing, B. and J. Poot, (2004): "Crossing boundaries and borders: Regional science advances in migration modelling", Papers in Regional Science, 83, 317-338.

[13] Davies, P. S., M. J. Greenwood, and H. Li (2001): "A Conditional Logit Approach to U.S. State-to-State Migration", Journal of Regional Science, 41, 337-360.

[14] Erickson, T. and T. Whited (2000) "Measurement Error and the Relationship between Investment and q", Journal of Political Economy, 108, 1027-57.

[15] Gourieroux, C., A. Monfort, and E. Renault (1993): "Indirect inference", Journal of Applied Econometrics, 8, S85-118.

[16] Greenwood, M. J. (1975): "Research on internal migration in the United States: A survey", Journal of Economic Literature, 13, 397-433.

[17] Greenwood, M. J. (1985): "Human migration: Theory, models and empirical studies", Journal of Regional Science, 25, 521-544.

[18] Greenwood, M.J. (1997): "Internal migration in developed countries", In: Rosenzweig, M. R. and O. Stark (eds.), Handbook of population and family economics, Volume 1B, Elsevier, Amsterdam.

[19] Hunt, G. L. and R. E. Mueller (2004): "North American Migration: Returns to Skill, Border Effects and Mobility Costs", Review of Economics and Statistics, 86, 988-1007.

[20] Kennan, J. and J. R. Walker (2003): "The Effect of Expected Income on Individual Migration Decisions", NBER Working Papers, No. 9585.

[21] Kennan, J. and J. R. Walker (2006): "The Effect of Expected Income on Individual Migration Decisions", mimeo, revised version, Universitiy of Wisconsin.

[22] Levin, A., C. F. Lin, C. S. J. Chu (2002): "Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties", Journal of Econometrics, 108, 1-24.

[23] Sjaastad, L. (1962): "The costs and returns of human migration", Journal of Political Economoy, 70, 80-93.

[24] Smith, T. (1993): "Estimating nonlinear time-series models using simulated vector autoregressions", Journal of Applied Econometrics, 8, S63-84.

[25] Storesletten, K., C. I. Telmer, and A. Yaronx (2004): "Cyclical Dynamics in Idiosyncratic Labor-Market Risk", Journal of Political Economy, 112, 695-717.

[26] Tauchen, G. (1986): "Finite state Markov-chain approximation to univariate and vector autoregressions", Economics Letters, 20, 177-81.

[27] Tunali, I. (2000): "Rationality of Migration", International Economic Review, 41, 893-920.