Schmollers Jahrbuch 129 (2009), 333 – 342
Duncker & Humblot, Berlin

Migration and Religion

Measuring Income Assimilation of Migrants to Germany

By Ira N. Gang, John Landon-Lane, and Myeong-Su Yun*

Abstract

We measure the income assimilation of migrants to Germany employing a new measure of assimilation that uses the whole income distribution rather than selected moments. To do this we implement a discrete-state Markov chain to model the dynamics of the cross-sectional income distribution of migrants and natives in Germany. Bayesian methods allow us to fully characterize the limiting cross-sectional income distribution for migrants and natives, enabling us to compare our measures of assimilation in the limiting case. We find no evidence in this sample of income assimilation for migrants to Germany.

JEL Classification: F22, J15, J16

1. Introduction

Assimilation of migrants in host countries has become a more watched phenomenon as international migration and globalization become more and more wide-spread. A large number of papers, including important early contributions from Chiswick (1978) and Borjas (1985), have studied the economic assimilation of migrants into host countries, typically asking whether and at what rate migrant wages (or some other measure) catch up to those of the native-born.

This paper examines how to measure assimilation in income. Indeed, assimilation can be defined from various aspects of the income (or other measure) distributions of natives and migrants. We develop a new measure of assimilation. Unlike other assimilation measures which are based on mean characteris-

* This paper was prepared for the SOEP2008 conference volume based on our presentation at the 8th International German Socio-Economic Panel User Conference, July 3 – 4, 2008, Berlin, Germany.

tics, ours captures assimilation across the entire distribution; testing whether two distributions of income of natives and natives are identical in the limit.

For our example we implement a discrete state Markov chain to model the dynamics of the cross-sectional income distribution of migrants and natives in Germany. We must note however, that our method can be applied to many other distributions such as, for example, occupation prestige. The discrete-state Markov chain model has been widely used for studying the dynamics of economic activities including, e.g., income and social mobility. We employ Bayesian methods in this paper which allow us to fully characterize the limiting cross-sectional wage distribution for migrants and natives, thus allowing us to compare our measures of assimilation in the limiting case. Finally we formally test the hypothesis of assimilation at the limiting distribution.

2. Method

In order to quantify the amount of assimilation present in a sample we need to be able to do two things: 1) we need to be able to compute the limiting behavior of a cross-sectional distribution and 2) formally compare the cross-sectional distributions for each population initially and then in the limit. If the difference between cross-sectional distributions significantly decline then we argue that this is evidence of assimilation. If the difference completely vanishes then we have full assimilation and if it declines then we have partial assimilation. First we describe how the limiting distribution is computed.

2.1 Computing the Limiting Income Distribution

In order to formally test for evidence of assimilation with respect to a variable of interest (in our case real income) we first need to be able to compute the limiting income distribution for both the native born and migrant populations. To do this we use a discrete state Markov chain to model the dynamics of the real income distribution. Discrete state Markov chains have a long history in the study of the dynamics of the income distribution starting with notable early contributions by Champernowne (1953) and Prais (1955). In using this model we first break the income distribution into discrete classes or bins. These income classes are mutually exclusive intervals that are defined to evenly cover the total income distribution. In fact we define the income classes so that they are equal in log length as first suggested by Champernowne (1953).

Let $\mathbb{D}$ be the range of values that real income can take. That is, assume $\mathbb{D} \subseteq \mathbb{R}^{+}$. Then define a set of disjoint intervals $\{C_1, \ldots, C_N\}$ such that $\mathbb{D} = C_1 \bigcup \cdots \bigcup C_N$. Let the vector $\pi_t = (\pi_{1t}, \ldots, \pi_{Nt})$ represent the (cross-sectional) probability distribution of income in period t. Letting y_{it} represent the income for an individual, *i,* in period t then π_{jt} is defined as $\pi_{jt} = prob\left(y_{it} \in C_j\right)$. That is, π_{jt} represents the unconditional probability that an individual's income falls in income class *j* in period *t.*

In order to test for assimilation we would first like to compare the cross-sectional income distribution, π_t, for both native born Germans and migrants to Germany. This initial comparison would set the benchmark for which we will subsequently be able to use in our test for evidence of assimilation. The definition of assimilation that we use is whether the two cross-sectional income distributions, for locals and migrants, appear to be converging to identical distributions. That is, over time does the cross-sectional income distribution of migrants to Germany converge to the same distribution that native born Germans converge to. To do this we need to compute the limiting distributions for each population and this requires that an assumption about the dynamics of the income distribution is made. The assumption that we make here is that the cross-sectional distribution of income, π_t, follows a first order Markov process.

That is we assume that

$$\pi_t^{'} = \pi_{t-1}^{'} P \quad \text{for} \quad t = 1, \ldots, \tag{1}$$

where the matrix P is an $(N \times N)$ probability transition matrix. The $(ij)^{th}$ element of P is the probability that an individual whose income in period $t-1$ is in income class C_i transitions to income class C_j by period *t.* Thus each row of P must sum to 1 as the set of income classes cover the whole space of possible incomes.

The dynamics of the income distribution over time is governed by the properties of *P.* In particular we are interested in the limiting distribution of the process that is given in (1). The limiting distribution, also referred to as the ergodic distribution, is the cross-sectional income probability distribution, $\bar{\pi}$, such that

$$\bar{\pi}^{'} = \bar{\pi}^{'} P\,. \tag{2}$$

The limiting cross-sectional income distribution implied by (1) is the left eigenvector of P associated with the eigenvalue 1.[1] Given an initial income distribution, π_0, then one can show that

$$\bar{\pi} = \lim_{t \to \infty} \pi_0^{'} P^t\,. \tag{3}$$

As discussed in Geweke (2005) the existence and uniqueness of $\bar{\pi}$ is governed by the eigenvalues of *P.* If there is only one eigenvalue of P that is equal

[1] See Geweke (2005) for a fuller discussion of the properties of the first order discrete state Markov model.

to 1 then the limiting distribution defined in (2) is unique and speed of convergence to $\bar{\pi}$ is related to the magnitude of the second highest eigenvalue of *P*. The closer the second highest eigenvalue of *P* is to 1, the slower the convergence to $\bar{\pi}$.

Now that we are able to compute the limiting cross-sectional distribution for any population we can now measure the amount of assimilation present in the data. This is described in the following subsection.

2.2 Measuring Assimilation

In order to measure the amount of assimilation between two populations we first need to measure the distance between the cross-sectional income distribution of two populations, *A* and *B*. If the initial cross-sectional distributions of population *A* and *B* are different, that is $\mu(\pi_0^A, \pi_0^B) \neq 0$, then we can look for evidence of assimilation. We do that by comparing the distance between the two population's limiting cross-sectional income distributions, $\mu(\pi_\infty^A, \pi_\infty^B)$, to the distance between the two population's initial income cross-sectional distribution, $\mu(\pi_0^A, \pi_0^B)$. We make the following definitions:

Definition 1: Assimilation in Distribution

Let $\mu(\pi_t^A, \pi_t^B)$ be a measure of distance between two discrete probability distributions such that

1. $\mu(\pi_t^A, \pi_t^B) = 0$ only if $\pi_{kt}^A = \pi_{kt}^B$ for all $k = 1, \ldots, N$.
2. $\mu(\pi_t^A, \pi_t^B) > 0$ if $\pi_{kt}^A \neq \pi_{kt}^B$ for some k.

Then we have *assimilation in distribution* between two populations if

1. $\mu(\pi_0^A, \pi_0^B) \neq 0$ and
2. $\mu(\pi_\infty^A, \pi_\infty^B) = 0$.

Our approach will be to estimate the initial distribution, π_0, for each population and then estimate the first order Markov model in (1). Using our estimates for the transition probability matrices, *P*, we then compute the limiting cross-sectional income distributions, π_∞, for each population. Then we test whether the two conditions in Definition 1 hold. To do this we need first to define our distance metric, $\mu(\pi^A, \pi^B)$, and second to estimate the parameters of interest and their standard errors.

There are many distance metrics we could use to define a distance between two discrete distributions but the one we use in this paper is the discrete version of the Kolmogorov-Smirnoff statistic. That is we define our distance metric to be

$$\mu\left(\pi^A, \pi^B\right) = \max_{i=1,\ldots,N} \left|\pi_i^A - \pi_i^B\right| . \tag{4}$$

This metric clearly satisfies the two conditions for our distance metric in Definition 1.

In order to test for evidence of assimilation we need to be able to test whether $\mu(\pi_0^A, \pi_0^B) \neq \mu(\pi_\infty^A, \pi_\infty^B)$ and $\mu(\pi_\infty^A, \pi_\infty^B) = 0$. These measures are highly non-linear functions of the parameters of our model given in (1). Thus computing standard errors and test statistics for the three main tests is going to very difficult using classical methods. However, Bayesian methods allow us to compute the exact finite sample posterior distributions for all parameters of the model given in (1) and of the statistics that we need in order to test for assimilation.

Bayesian methods were first used to estimate models such as (1) by Geweke et al. (1986). For the sake of brevity we do not discuss the method used to estimate (1) in detail.[2] A good summary of the Bayesian methods used to estimate (1) can be found in Geweke (2005). What we do in this paper is to use conjugate priors for π_0 and P in (1) so that we obtain posterior distributions for these parameters that are known and can be drawn from directly. The initial distribution, π_0, and each row of P are multivariate Beta distributions so if we use multivariate Beta prior densities then the posteriors are also multivariate Beta. Independent and identically distributed (i.i.d) samples can then be obtained from the posterior distribution of (1) using the method described in Devroye (1986).

Once we have i.i.d samples from the posterior distribution for π_0 and P we can then construct i.i.d samples for π_∞ and for the various distance metrics we need. We can then look at the posterior distribution of $\mu(\pi_0^A, \pi_0^B)$, $\mu(\pi_\infty^A, \pi_\infty^B)$, and $\mu(\pi_\infty^A, \pi_\infty^B) - \mu(\pi_0^A, \pi_0^B)$ and compute standard errors and highest posterior density intervals for each statistic.

3. Testing for Assimilation of German Migrants

In this section we describe the data that we used in this study and report the results of our measures of income assimilation for migrants to Germany.

3.1 Data

In this paper we use data from the German Socio-Economic Panel (SOEP) for the year from 1985 until 2005. In order to observe meaningful transitions in the income distribution we define 5 year transitions for our study. That is, the time span between period t and period $t+1$ in this study is 5 years. We concentrate our attention to individuals between the ages of 25 and 55 at the

[2] A full technical appendix describing the Bayesian methods used can be obtained from the authors upon request.

start of each transition. We do this so as to restrict our attention to those individuals who are not yet retired and also not likely to retire during the period and those who most likely left formal studies before the start of the period.

We use SOEP Samples A and B. At the beginning of the SOEP in 1984, it was the head of the household's nationality which defined the classification in both samples A and B. The identity of sample B is often assumed to be that of the population of the group of "foreigners" surveyed by the SOEP, while sample A contains "Germans". For the most part this is correct, although it is not precise and over time becomes less accurate. First, it is possible that there are other household members present with a different nationality to that of the head of the household. In addition, sample A contains foreigners whose nationality was not represented in sample B. In 1984 1.7 % (149 / 8927) of Sample A were Non-German; this changes to 1.4 % (73 / 5077) in 2005. In 1984 3.9 % (120 / 3049) of Sample B were Germans; becoming 39.96 % (406 / 1016) in 2005. The increase in Germans in Sample B over time will cause the two samples to become more similar which would bias our results towards finding assimilation.

Table 1

Income class definitions (monthly income; constant 2001 Euros)

Income Class	Lower Bound	Upper Bound
1	0	364.23
2	364.23	678.04
3	678.04	1,262.06
4	1,262.06	2,349.13
5	2,349.13	∞

3.2 Definition of Income Classes and Priors

To test for income assimilation we use monthly net labor income in 2001 constant Euros. The income classes are chosen to be equal in log length and to evenly cover the aggregate distribution of net monthly income over the whole sample period. That is, the income classes are defined over the monthly incomes for all individuals for all the years in our sample. Table 1 contains the income class definitions. There are 5 income classes in total and these classes range from very low income (class 1) to very high income (class 5). These income classes are designed so that we do not have the problem of the distribution "piling" up at one extreme or the other. To do this we define the highest and lowest class to contain the highest 5 % and lowest 5 % of the combined income distribution respectively. The intervening classes are defined to have equal log-length as suggested by Champernowne (1953).

There are two parameters that make up the model that is estimated in (1). First there is the initial distribution, π_0, and second there is the probability transition matrix, P. The priors used for each parameter are as follows: For the initial distribution we use a multivariate-Beta distribution that is indexed by the vector $a_0 = (2, 2, 2, 2, 2)'$. This implies a "flat" prior for π_0 with prior mean for π_{0i} of 0.2 for $i = 1, \ldots, 5$ and a prior standard deviation for π_{0i} of 0.35 for $i = 1, \ldots, 5$. The prior for P is similar in that for each row of P we use a multivariate-Beta prior. Thus the prior for P is the product of independent multivariate-Beta priors indexed by the appropriate row of A where

$$A = \begin{bmatrix} 9.9 & 0.025 & 0.025 & 0.025 & 0.025 \\ 0.025 & 9.9 & 0.025 & 0.025 & 0.025 \\ 0.025 & 0.025 & 9.9 & 0.025 & 0.025 \\ 0.025 & 0.025 & 0.025 & 9.9 & 0.025 \\ 0.025 & 0.025 & 0.025 & 0.025 & 9.9 \end{bmatrix}. \tag{6}$$

This prior loads most of the prior probability onto the diagonal of P which yields a process by which there is very little mobility. The implied limiting distribution under this prior is $\pi_\infty = (0.2, 0.2, 0.2, 0.2, 0.2)'$. Thus the prior is a "no-change" prior. The prior is also defined so that it is relatively diffuse so that the posterior distribution reflects mainly information from the sample and not from the prior.[3]

3.3 Results

Using the priors defined in Section 3.2 we estimate (1) using the method described in Geweke (2005). Most importantly we compute posterior distributions for $\mu(\pi_0^A, \pi_0^B)$, $\mu(\pi_\infty^A, \pi_\infty^B)$, and $\mu(\pi_\infty^A, \pi_\infty^B) - \mu(\pi_0^A, \pi_0^B)$. We compute these distributions for each of the four 5-year transitions; 1985 – 1990, 1990 – 1995, 1995 – 2000, and 2000 – 2005. We estimate (1) for each transition separately rather than pooling the four transitions together to allow for P to vary over the whole sample. Thus, we can check whether the evidence of assimilation (if any) changes over time. Table 2 reports the posterior moments for the three measures described above. In particular it reports the posterior mean, posterior standard deviation (in parentheses), and 95 % highest posterior density interval (in brackets) for each of the three measures for each 5-year transition.

The overall results reported in Table 2 do not show any evidence of assimilation for this sample. When we look at the gap between the limiting distributions implied by the estimation transition probability matrices, P, for each tran-

[3] We cannot define a minimum entropy prior here as this would imply 0 off diagonal probability in the prior. For practical purposes we need some non-zero prior probability placed on the off-diagonal elements of P so the prior chosen is as close to a minimum entropy prior as practically possible.

sition we see that this gap is increasing and not decreasing. Hence we see no formal evidence of income assimilation in this sample for any of the four transition periods. When we look at the limiting distributions[4] we see that the German born (Sample A) income distributions shifts more to the right than migrants. In fact we see that the income class where that attains the maximum difference is always the highest income class in the limiting distribution suggesting that migrants are being affected by some form of "glass ceiling" and / or "sticky floor" in their income mobility.

Table 2

Measures of assimilation: ages 25 – 55

Transition	$\mu\left(\pi_0^A, \pi_0^B\right)$	$\mu\left(\pi_\infty^A, \pi_\infty^B\right)$	$\mu(\pi_\infty^A, \pi_\infty^B) - \mu(\pi_0^A, \pi_0^B)$
1985 – 1990	0.225 (0.017) [0.192, 0.259]	0.376 (0.049) [0.268, 0.466]	0.151 (0.052) [0.040, 0.250]
1990 – 1995	0.115 (0.009) [0.099, 0.132]	0.204 (0.034) [0.125, 0.265]	0.090 (0.035) [0.012, 0.154]
1995 – 2000	0.114 (0.011) [0.094, 0.137]	0.200 (0.056) [0.082, 0.304]	0.085 (0.057) [–0.036, 0.188]
2000 – 2005	0.110 (0.012) [0.088, 0.134]	0.267 (0.068) [0.123, 0.390]	0.157 (0.069) [0.009, 0.280]

One point to note about the results reported in Table 2 is that evidence of assimilation may be hidden by the mixing of all ages in each year. It could be that younger migrants are much more similar to younger Germans than is the case for older age groups. Because all ages are mixed together in Table 2 we may not be picking this up. Table 3 reports the differences between the cross-sectional income difference in each transition for a number of different cohorts. Each cell of Table 3 reports the posterior mean and standard deviation of $\mu\left(\pi_0^A, \pi_0^B\right)$. Each column of Table 3 reports results for a distinct cohort. The cohort is defined by the year in which they were in the 25 to 30 year old age group and we can follow the behavior of each age group by looking along each diagonal (from left to right).

For each year (i.e. along each row of Table 3) we observe that the gap between Germans and migrants is larger for older age groups than for younger age groups and we see this as well for each cohort (each column of Table 3).

4 Not reported here but available from the authors upon request.

That is, as each cohort gets older the gap between Germans and migrants does not disappear and in some cases getting larger. Looking down each diagonal we see that there is some reduction in the gap between Germans and migrants over time but this is only significant from 1985 to 1990. After 1990 there does not appear to be any further decrease of the gap between income distributions. These results reinforce the results reported in Table 2 that there is no evidence of income assimilation for migrants to Germany for this sample.

Table 3

Comparison of initial income distributions, $\mu\ (\pi_0^A, \pi_0^B)$, by cohort year in which cohort was aged 25 – 30 years

Year	1965	1970	1975	1980	1985	1990	1995	2000
1985	0.27 (0.04)	0.29 (0.04)	0.30 (0.04)	0.13 (0.02)	0.08 (0.03)			
1990	0.18 (0.03)	0.21 (0.03)	0.20 (0.03)	0.09 (0.03)	0.09 (0.03)	0.11 (0.01)		
1995		0.22 (0.05)	0.22 (0.03)	0.22 (0.06)	0.15 (0.04)	0.07 (0.03)	0.06 (0.03)	
2000			0.23 (0.04)	0.13 (0.04)	0.17 (0.03)	0.08 (0.03)	0.07 (0.03)	0.07 (0.03)

4. Conclusion and Discussion

In this paper we investigated whether there was any evidence for assimilation in income for migrants to Germany using data from the German Socio Economic Panel. The measure of assimilation used all information from the income distribution and tested whether the observed dynamics implied whether the limiting income distributions of Germans and migrants to Germany were getting closer together. We found no evidence of this and, if anything, we found that the gap was getting bigger. We found that the differences in the distributions between migrants and natives was most pronounced at the higher incomes.

It should be noted that the results reported here are for those migrants who stayed in Germany after arrival. These are the "successful" migrants who did not leave Germany in the period between arriving (prior to 1974) and when the sample was first taken (1984). We do not explicitly model the reasons why individuals, both natives and migrants, leave the sample as we do not have information on why some individuals disappear from the panel. Given our results that the lower incomes appear to be more assimilated than the higher incomes then we don't think our results are affected by attrition caused by individuals leaving the labour force. The issue of migrants leaving and returning is left for future research but it must be remembered that the migrants in our sample have already been in Germany for over ten years before they were sampled. In our view these migrants have a much more permanent attachment to Germany than more recent migrants would have so that we believe our results are not as affected by the issue of return migration as would be the case if we studied new migrants.

References

Borjas, G. (1985): Assimilation, change in cohort quality, and the earnings of immigrants, Journal of Labor Economics 3 (4), 463 – 489.

Champernowne, D. G. (1953): A model of income distribution, Economic Journal 63 (250), 318 – 351.

Chiswick, B. (1978): The effect of Americanization on the earnings of foreign-born men, Journal of Political Economy 86 (5), 897 – 921.

Devroye, L. (1986): Non-uniform Random Variate Generation, New York.

Geweke, J. F. (2005): Contemporary Bayesian Econometrics and Statistics, Wiley Series in Probability and Statistics, Wiley-Interscience, Hoboken, New Jersey.

Geweke, J. F. / *Marshall,* C. / *Zarkin,* G. A. (1986): Exact inference for continuous time markov chain models, Review of Economic Studies 53 (4), 653 – 669.

Prais, S. (1955): Measuring social mobility, Journal of the Royal Statistical Society, Series A Part I (118), 56 – 66.