Second Order Approximation for the Average Marginal Effect of Heckman's Two Step Procedure

Alpaslan Akay*, Elias Tsakas†

University of Göteborg, Department of Economics, P.O.Box 640, 405 30 Göteborg, Sweden.

January, 2006

Abstract

In this paper we discuss the differences between the average marginal effect and the marginal effect of the average individual in sample selection models, estimated by Heckman's two step procedure. We show that the bias that emerges as a consequence of interchanging them, could be very significant, even in the limit. We suggest a computationally cheap approximation method, which corrects the bias in a large extent. We illustrate the implications of our method with an empirical application of earnings assimilation and a small Monte Carlo simulation.

Keywords: *Heckman's two step estimator, average marginal effect, marginal effect of the average individual, earnings assimilation.*

JEL Classification: C13, C15, J40.

*Tel:+46-31-773 5304; Fax:+46-31-773 4154; E-mail: Alpaslan.Akay@economics.gu.se

†Tel:+46-31-773 4106; Fax:+46-31-773 4154; E-mail: Elias.Tsakas@economics.gu.se

1 Introduction

The widely established point estimators for the marginal effect of an explanatory variable in a parametric model are (i) the sample average marginal effect and (ii) the marginal effect of the sample average individual. In general, neglecting their quantitative and, more importantly, conceptual differences is a quite common practice, which could lead to misleading results, especially in non-linear models. Arbitrarily interchanging them could create systematic bias since the two measures estimate different quantities.

In this paper we study the average marginal effects of Heckman's very popular two step estimation procedure, which is widely used, especially in labor supply studies. Provided that one is interested in the average effect over the population, rather the effect over the average individual, we show that evaluating the derivative at the sample means, leads to biased predictions, even asymptotically. Since the other commonly used alternative (averaging the marginal effects for the whole sample), though consistent, could be computationally inefficient, we propose an approximation technique which significantly reduces the bias, without increasing much the number of numerical operations. In order to so, we approximate the average marginal effect with a Taylor expansion and prove that the conventionally used marginal effect of the average individual is actually equal to the first order Taylor approximation, while the order of magnitude is equal to the asymptotic bias. By shifting to the second order approximation, one can reduce the size of the bias without high computational cost, since the second term of the series is a function of the Hessian and the covariance matrix evaluated at the sample means.

In order to emphasize the necessity of a consistent estimator for the average marginal effects, we present an empirical application of immigrants earnings assimilation using registered data from Sweden. We find that our approach corrects the bias in a very large extent and we discuss the policy implications behind this relative difference.

The paper has the following structure. Section 2 briefly describes Heckman's two step procedure. In section 3 we introduce the theoretical results of our approach. In section 4 we apply the model to real data. In section 5 we include Monte Carlo simulations. Section 6 concludes.

2 Heckman's two step procedure and marginal effects

Consider the following sample selection model

$$\left\{ \begin{array}{l} Y_i^* = \mathbf{X}_i'\beta + \epsilon_i \\ H_i^* = \mathbf{Z}_i'\gamma + u_i \\ H_i = 1[H_i^* > 0] \end{array} \right\} \tag{1}$$

where $i = 1, ..., N$. Let the latent variables Y_i^* and H_i^* denote individual i's earnings and hours of work respectively. Assume also that the matrices $\mathbf{X}_i$ and $\mathbf{Z}_i$ include various observed individual characteristics, with $\mathbf{X}_i$ being a strict subset of $\mathbf{Z}_i$. Finally the joint error term (ϵ_i, u_i) follows bivariate normal distribution with correlation coefficient ρ. Our primary aim is to estimate the parameter vector β of the earnings equation. However neither Y_i^*, nor H_i^* are observed. On the other hand, we know that strictly positive hours of work is necessary and sufficient condition for participating in the job market , ie. $H_i^* > 0$. Then the participation decision takes the form of a binary choice[1], since *working* and *not working* are complementary events, and as such they can be written as the indicator function of equation (1).

Conditioning on the subset of the population that contains the individuals who actually work, the expectation of the earnings given participation would be given by the following formula:

$$\begin{aligned} E[Y_i^*|H_i = 1, \mathbf{X}_i, \mathbf{Z}_i] &= E[\mathbf{X}_i'\beta + \epsilon_i | H_i^* > 0] \\ &= \mathbf{X}_i'\beta + E[\epsilon_i | u_i > -\mathbf{Z}_i'\gamma] \\ &= \mathbf{X}_i'\beta + \rho\sigma_\epsilon \frac{\phi(-\mathbf{Z}_i'\gamma/\sigma_u)}{1 - \Phi(-\mathbf{Z}_i'\gamma/\sigma_u)} \end{aligned} \tag{2}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ denote the density and the cumulative distribution of a standard normal distribution respectively. After some notation simplification equation (2) is rewritten as follows:

$$E[Y_i^*|H_i = 1, \mathbf{X}_i, \mathbf{Z}_i] = \mathbf{X}_i'\beta + \rho\sigma_\epsilon\lambda(\alpha_u) \tag{3}$$

[1] The model can be extended to multinomial discrete choice framework.

where $\alpha_u = -\mathbf{Z}_i'\gamma/\sigma_u$, while λ denotes the inverse of the Mill's ratio, ie. $\lambda = \phi/(1-\Phi)$. It is straightforward that equation (3) cannot be estimated consistently with ordinary least squares (OLS) in the existence of correlation between ϵ_i and u_i ($\rho \neq 0$). On the other hand, although consistent, the maximum likelihood estimator (MLE) constitutes a computationally challenging task. Heckman (1976) introduced a method which can simultaneously handle consistency and computational efficiency. His procedure consists of two separate steps. First estimate the participation probability by applying a binary probit model

$$P[H_i = 1|\mathbf{Z}_i] = \Phi(\mathbf{Z}_i'\gamma) \tag{4}$$

and use the estimated choice probabilities to calculate $\lambda(\alpha_u)$. In the second step, apply OLS on the earnings equation, while perceiving the estimated inverse Mill's ratio as another explanatory variable. Thus one gets rid of the omitted variable problem that would emerge otherwise and the estimator of the parameter vector β becomes consistent.

The *ceteris paribus* marginal effect[2] of an infinitesimal change of an arbitrary individual characteristic k on individual i's earnings is given by the following equation for an explanatory variable $x_{k,i}$

$$ME_{k,i} = \frac{\partial E[Y_i^*|H_i = 1, \mathbf{X}_i, \mathbf{Z}_i]}{\partial X_{k,i}} = \beta_k - \gamma_k \rho \sigma_\epsilon \delta(\alpha_u) \tag{5}$$

where $\delta(\alpha_u) = \lambda^2(\alpha_u) - \alpha_u \lambda(\alpha_u)$. Notice that the previous expression can be decomposed into two distinct terms, a constant and a dependent on the explanatory variables[3] one. In the relevant literature this separation is quoted as *distinction between direct and indirect effect*. The total marginal effect of a variable that is included in $\mathbf{Z}_i$ but not in $\mathbf{X}_i$ would be equal to the indirect effect, since the direct effect in this case would be equal to 0. Henceforth, for notation simplicity and without loss of generality, unless stated differently we will omit $\mathbf{X}_i$ and $\mathbf{Z}_i$ from the conditional expectations.

[2] A more precise terminology would require to define it as *conditional marginal effect*, since it refers only to the individuals who actually work.

[3] Since $\mathbf{X}_i$ is a strict subset of $\mathbf{Z}_i$ the explanatory variables of the earnings equation are included in the second term.

3 Average marginal effect decomposition

One striking feature of non-linear models is that, due to the functional relationship between individual characteristics and the derivative of the earnings equation, marginal effects vary across individuals. Thus when policy makers decide upon an action which leads to a change of an explanatory variable which affects the whole population, they usually take into account the average marginal effect (AME) across all individuals. Therefore using an inconsistent estimator for AME could potentially lead to wrong conclusions and undesired effects of the policy application. The aim of this section is to clearly show that the marginal effect of the sample average individual ($\widehat{MEAI}$), is not only biased for small samples, but also inconsistent estimator of the AME. We suggest that one should use the sample's average marginal effect ($\widehat{AME}$), which is a consistent estimator for the AME.

More precisely, for the sample selection model of equation (1), the $\widehat{AME}$ is given by the following equation for an explanatory variable X_k

$$\widehat{AME}_k = \frac{1}{N}\sum_{i=1}^{N}\frac{\partial E[Y_i^*|H_i=1]}{\partial X_{k,i}} = \frac{1}{N}\sum_{i=1}^{N}\big(\beta_k - \gamma_k\rho\sigma_\epsilon\delta(\alpha_u)\big) \tag{6}$$

It follows directly from Khinchine's weak law of large numbers that $\widehat{AME}$ is a consistent estimator of the AME. Namely

$$\text{plim}_{N\to\infty}\widehat{AME}_k = E_{\mathbf{Z}}[\beta_k - \gamma_k\rho\sigma_\epsilon\delta(\alpha_u)] \tag{7}$$

for every k. On the other hand the corresponding estimated marginal effect of the average individual is given by the following formula.

$$\widehat{MEAI}_k = \left.\frac{\partial E[Y_i^*|H_i=1]}{\partial X_{k,i}}\right|_{\mathbf{Z}_i=\overline{\mathbf{Z}}} \tag{8}$$

It is straightforward then that

$$\text{plim}_{N\to\infty}\widehat{MEAI}_k = \beta_k - \gamma_k\rho\frac{\sigma_\epsilon}{\sigma_u}\delta(\mathbf{M}'\gamma) \tag{9}$$

where $\mathbf{M}$ denotes the vector of the expected values of $\mathbf{Z}_i$ which is constant across individuals. In order to extract the asymptotic bias of $\widehat{MEAI}_k$, which appears due to non-linear derivatives, we expand the Taylor series[4] of $\delta(\mathbf{Z}_i'\gamma)$

[4] Verlinda (forthcoming) applies a similar method on the binomial probit model.

around $\mathbf{M}$. That is,

$$\delta(\mathbf{Z}_i'\gamma) = \delta(\mathbf{M}'\gamma)+\sum_{j=1}^{\infty}\left[\frac{1}{j!}\sum_{k_1,\dots,k_j}\left(\left.\frac{\partial^j\delta(\mathbf{Z}_i'\gamma)}{\partial Z_{k_1,i},\dots,\partial Z_{k_j,i}}\right|_{\mathbf{M}}\cdot(Z_{k_1,i}-M_{k_1})\cdots(Z_{k_j,i}-M_{k_j})\right)\right] \tag{10}$$

Then plugging into equation (5) and taking expectation we conclude that, since consistency holds, the AME is approximated by the following formula

$$\begin{aligned}\text{plim}\widehat{AME}_k &= \beta_k - \gamma_k\rho\frac{\sigma_\epsilon}{\sigma_u}E_{\mathbf{Z}}[\delta(\mathbf{Z}_i'\gamma)] \\ &= \text{plim}\widehat{MEAI}_k - \gamma_k\rho\frac{\sigma_\epsilon}{\sigma_u}\sum_{j=1}^{\infty}\left[\frac{1}{j!}\sum_{k_1,\dots,k_j}\left(\left.\frac{\partial^j\delta(\mathbf{Z}_i'\gamma)}{\partial Z_{k_1,i},\dots,\partial Z_{k_j,i}}\right|_{\mathbf{M}}\cdot\Psi^j_{k_1,\dots,k_j}\right)\right] \\ &= \text{plim}\widehat{MEAI}_k - B_k^1(\Psi^1,\Psi^2,\dots) \end{aligned} \tag{11}$$

where $\Psi^j_{k_1,\dots,k_j} = E_{\mathbf{Z}}[(Z_{k_1,i} - M_{k_1})\cdots(Z_{k_j,i} - M_{k_j})]$ denotes the j^{th} order joint moment about the means, while B_k^1 denotes the size of the first order approximation asymptotic bias as a function of the joint moments, Ψ^j, of the individual characteristics. Therefore by using the $\widehat{MEAI}_k$ to estimate the AME_k one implicitly takes into account only the first order approximation while neglecting the higher orders. By using the second order approximation, which does not increase significantly the number of numerical operations since it only involves the Hessian evaluated at $\mathbf{M}$ and the covariance matrix, one would reduce[5] the bias to B_k^2.

In the following section we empirically show that neglecting the bias could create misleading results that could significantly affect the policy implications of the model.

4 Empirical application: an economic assimilation study

In order to illustrate the importance of the previous analysis we provide an application from earnings assimilation theory[6]. The central question in such

[5] The expected second order of magnitude is larger to the third one (Nguyen, Jordan; 2004).

[6] Selection in such studies can arise either as self-selection by the individuals or as sample selection by the data analyst. Not taking it into consideration could significantly distort the inference.

a study would typically be *if and when the earnings of the immigrants catch up with the ones of the native population* (Borjas, 1985, 1999; Longva *et.al.*, 2003). Then, based on the answer, policies that target to different individual characteristics of the immigrants are designed, in order to adjust the speed of assimilation closer to the desired for the policy maker level. Thus estimating consistently the average marginal effects is rather crucial.

The data used for the purpose of the present study comes from the registered nationally representative longitudinal individual data set of Sweden (LINDA). The working sample includes 3136, aged 18-65, male individuals (1962 immigrants[7] and 1174 natives) observed for 11 years. Table 4 shows the mean characteristics of the sample.

The model specification for the immigrants is given by equation (1). Keeping track with most similar studies, we use the natural logarithm of the disposable income in the earnings equation. The individual characteristics included in X_i matrix are individual i's age, squared age, years since migration, years since migration squared, number of children and the dummies for marital status, size of the permanent residence area, education, arrival cohort and geographical origin. Since the time period cannot be identified together with the arrival cohort and the age we assume that its effect is the same for both natives and immigrants (Borjas, 1985, 1999). The Z_i matrix includes the same characteristics plus[8] the logarithm of other income. In the case of the natives we exclude the variables that do not make sense, such as years since immigration, arrival cohort and geographical origin.

The estimation results and the bias analysis for the probit equation (first step) and the target equation (second step) are presented in tables 5 and 6 respectively. A really interesting, though not surprising given the structure of the Taylor series, result is that the percentage change of the bias level by shifting to the second order approximation remains constant across explanatory variables. Notice that, as expected, the second order (SO) bias is significantly smaller than the first order (FO) one for both the immigrants and the natives

[7]We define an immigrant as an individuals who was born abroad (first generation).

[8]The set of the explanatory variables included in the earnings equation must be a strict subset of the corresponding set of the participation equation.

and both equations. Table 1 shows the size of the relative improvement if the second order approximation is used.

- *Table 1 about here*

Taking a closer look at the first and second order bias estimates of the earnings equation (table 6), one could easily notice the rather significant improvement, not only in relative, but also in absolute terms, especially in key variables, such as geographical origin or years since migration for the immigrants and age for the natives. Also in the selection equation, one can observe a quite large reduction in the absolute level of bias, especially for the immigrants. This becomes even more worth mentioning, since it is observed in key variables, such as other income. Additionally there is a quite remarkable absolute reduction in the bias of the constant parameter for both groups.

A natural question that rises at this point is *how important the difference between first and second order bias could be, especially in terms of policy implications*. The following example shows that there is a rather significant difference indeed. In the literature of earnings assimilation the marginal assimilation rate of immigrant i with respect to native j is given by the following formula:

$$MRA_{i,j} = ME_{\text{AGE},i} + ME_{\text{YSI},i} - ME_{\text{AGE},j} \tag{12}$$

However, what is really interesting for the policy maker is whether the average earnings of the immigrants, $E[Y_I]$, catch up with the average earnings of the natives, $E[Y_N]$. In order to estimate the average marginal rate of assimilation ($AMRA$) consistently, one should integrate over all combinations of natives and immigrants[9]. Namely,

$$\begin{aligned} \widehat{AMRA} &= \sum_{i=1}^{I}\sum_{j=1}^{N}\frac{1}{I}\frac{1}{N}\left(ME_{\text{AGE},i} + ME_{\text{YSI},i} - ME_{\text{AGE},j}\right) \\ &= \frac{1}{I}\sum_{i=1}^{I} ME_{\text{AGE},i} + \frac{1}{I}\sum_{i=1}^{I} ME_{\text{YSI},i} - \frac{1}{N}\sum_{j=1}^{N} ME_{\text{AGE},j} \\ &= \widehat{AME}^{I}_{\text{AGE}} + \widehat{AME}^{I}_{\text{YSI}} - \widehat{AME}^{N}_{\text{AGE}} \end{aligned} \tag{13}$$

[9]Unlike the marginal effect which is defined for one individual, the marginal rate of assimilation is defined for a pair of individuals.

Similarly the marginal rate of assimilation of the sample average individual ($MRAAI$), would be equal to

$$\widehat{MRAAI} = \widehat{MEAI}^{I}_{\text{AGE}} + \widehat{MEAI}^{I}_{\text{YSI}} - \widehat{MEAI}^{N}_{\text{AGE}} \tag{14}$$

and would estimate the $AMRA$ inconsistently, due to the bias in $MEAI$. Therefore by using the first order approximation of the AME, one would implicitly estimate the length of the assimilation period for the average individual ($LAPAI$), instead of the average length of assimilation ($ALAP$) which is what we would like to estimate in the first place. In such a case it is clear that the estimator would be inconsistent.

Table 2 shows the initial earnings difference and the length of the assimilation period, together with the first and second order bias, for immigrants from every geographical origin. These results are not surprising at all, since they clearly reveal faster assimilation rates for immigrants coming countries with higher average human capital level compared to those coming from developing ones. In any case, the really interesting point for our analysis is that by using the second order approximation we manage to reduce the bias for every single group up to approximately 1 year.

- *Table 2 about here*

5 Monte Carlo simulation

As we have already discussed the bias that emerges by using the $\widehat{MEAI}$ as a point estimator of the AME, is not a consequence of a small sample, which would disappear in the limit. Regardless of the sample size, higher order approximations lead to bias reduction compared to lower ones. The purpose of this section is to provide empirical evidence through a Monte Carlo experiment.

- *Table 3 about here*

Assume a classical sample selection model of the form of equation (1) with $\mathbf{X}_i$ being a singleton and $\mathbf{Z}_i = (Z_{1,i}, Z_{2,i})$ coming from a bivariate normal distribution with mean $\mu_i = (\mu_1, \mu_2)$ and covariance matrix Σ. Assume also the following parameter values $\beta = 1$, $\gamma = (3, -2)$, $\sigma_\epsilon = 0.5$, $\sigma_u = 1$, $\rho = -0.8$ and the following random generating process $\mu = (0.5, 1.5)$, $\Sigma = \left[\begin{smallmatrix} 0.5 & -0.1 \\ -0.1 & 1 \end{smallmatrix}\right]$. Then using pseudo-random numbers we repeatedly evaluate the first and the second order bias, while increasing the sample size with step of 100 observations. The results are presented on table 3.

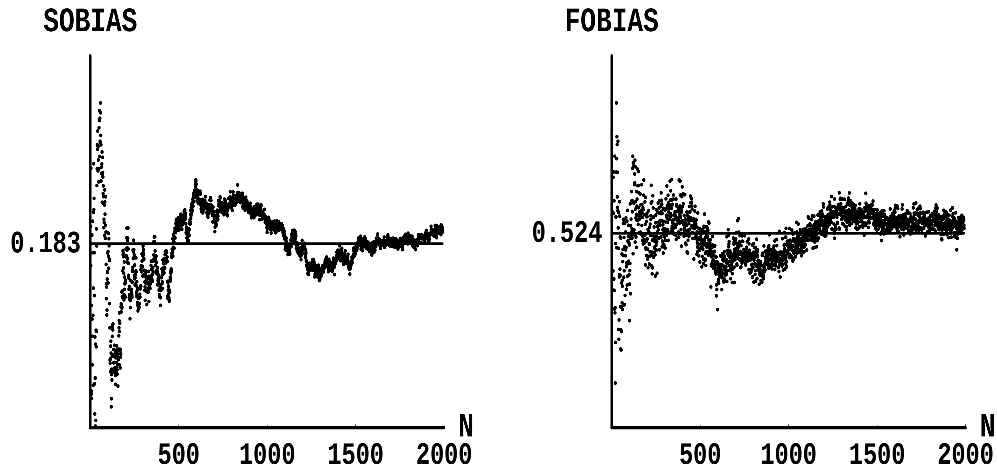

Figure 1: First and second order bias in Monte Carlo experiment.

Figure 1 illustrates the same point as table 3. Namely, it becomes clear that the bias that emerges by using the $MEAI$, is corrected in a rather large extent, without a corresponding computational cost. Notice that, bias reduction is observed, not only for small samples, but asymptotically too.

6 Concluding discussion

In this paper we discuss the differences between two point estimators of the marginal effect of an explanatory variable on the population, in a sample selection model estimated by Heckman's two step procedure. We show that on the contrary to a rather widespread perception that neglects possible differences between them, the average marginal effect is significantly different from the marginal effect of the average individual even asymptotically. Thus, it should be clear that there is not only a quantitative distinction, but also a conceptual one between these measures. Given that the usual aim is to extract information about the average effects on the population, a clear bias would emerge if using the marginal effect of the sample average individual. Hence we suggest an approximation method, based on Taylor expansion, which would correct the bias in a rather remarkable extent, while increasing relatively little the number of computational operations. Such an example is presented in the paper, alongside with a Monte Carlo experiment, and they both support the previous argument. Before closing, we would like to make clear that we do not argue in favor of the average marginal effect and against the marginal effect of the average individual. Our aim is to stress that, once the average marginal effect has been chosen as an informative tool for policy making, the sample marginal effect of the average individual provides inconsistent estimations which can be corrected in a large extent by the proposed method.

Acknowledgements

We are indebted to Lennart Flood, Marcela Ibanez, Florin Maican and Kerem Tezic for their valuable comments.

References

Borjas G.J. (1985). "Assimilation, changes in cohort quality, and the earnings of immigrants", *Journal of Labor Economics*, 3, 463-489.

Borjas G.J. (1999). "The economic analysis of immigration", *Handbook of Labor Economics*, edited by O. Ashenfelter and D. Card, vol. 3A, 1697-1760.

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

Greene W.H. (2003). "*Econometric analysis*", 5^{th} edition, Prentice Hall, Saddle River.

Heckman J.J. (1972). "The common structure of statistical models of truncation, sample selection and limited dependent variables, and a simple estimator for such models", *Annals of Economic and Social Measurement*, 5, 475-492.

Heckman J.J. (1979). "Sample selection bias as a specification error", *Econo-*

metrica, 47, 153-162.

Longva P., Raaum O. (2003). "Earnings assimilation of immigrants in Norway-a reappraisal", *Journal of Population Economics*, 16, 177-193.

Nguyen X., Jordan M.I. (2004). "On the concentration of expectation and approximate inference in layered networks", *Advances in Neural Information Processing Systems (NIPS)*, 16.

Saha A., Capps O.Jr., Byrne P.J. (1997). "Calculating marginal effects in models for zero expenditures in household budgets using Heckman-type correction", *Applied Economics*, 29, 1311-1316.

Saha A., Capps O.Jr., Byrne P.J. (1997). "Calculating marginal effects in dichotomous-continuous models", *Applied Economics Letters*, 4, 181-185.

Verlinda J.A. (2005). "A comparison of two common approaches for estimating marginal effects in binary choice models", forthcoming in *Applied Economics Letters*.

Appendix

Table 1: Mean characteristics of immigrants and natives.

	Immigrants		Natives	
Variables	Mean	St. Deviation	Mean	St. Deviation
Log earnings	8.5707	5.2519	10.7750	3.7428
Log other income	0.5656	1.9748	0.7746	2.3281
Age	37.14	11.03	38.37	11.27
Age squared	1501.0	866.0	1599.0	907.0
Big city ($> 250,000$)	0.6347	0.4815	0.7349	0.4414
Number of children	0.4840	0.9875	0.4407	0.8959
Married/Cohabiting	0.4344	0.4957	0.3891	0.4876
YSM	12.66	8.64	-	-
YSM squared	2348.6	277.1	-	-
Education (highest level):				
Upper-secondary	0.4454	0.4970	0.4867	0.4998
University	0.2591	0.4381	0.2744	0.4462
Arrival cohort:				
1970-1974	0.0693	0.2539	-	-
1975-1979	0.0927	0.2900	-	-
1980-1984	0.0807	0.2723	-	-
1984-1989	0.1499	0.3569	-	-
1990-1994	0.1589	0.3655	-	-
1995-2000	0.0358	0.1857	-	-
Geographical origin:				
Nordic	0.0968	0.2957	-	-
W. Europe (incl. EU)	0.0383	0.1918	-	-
USA	0.0443	0.2058	-	-
Eastern Europe	0.0799	0.2711	-	-
Middle East	0.0903	0.2866	-	-
Asia	0.0846	0.2783	-	-
Africa	0.0855	0.2797	-	-
Latin America	0.1075	0.3097	-	-

Table 2: Relative reduction of the bias.

	Immigrants	Natives
Participation equation	0.732	0.178
Earnings equation	0.735	0.943

Table 3: Estimates and analysis of bias for the assimilation period.

Variables	ALAP	FO LAPAI	SO LAPAI	FO Bias	SO Bias
Nordic	13.6973	12.7850	13.1966	0.9123	0.5006
W. Europe (incl. EU)	8.6961	8.1169	8.3782	0.5792	0.3178
USA	8.9012	8.3083	8.5758	0.5929	0.3253
Eastern Europe	15.4322	14.4043	14.8682	1.0279	0.5641
Middle East	23.9514	22.3561	23.0760	1.5953	0.8754
Asia	20.8989	19.5069	20.1351	1.3920	0.7639
Africa	25.3264	23.6395	24.4007	1.6869	0.9256
Latin America	19.0115	17.7452	18.3166	1.2663	0.6949
Total	16.9894	15.8578	16.3684	1.1316	0.6210

Note: Average length of the assimilation period (ALAP), first (FO) and second order (SO) approximation of the length of the assimilation period of the average individual (LAPAI) and first (FO) and second (SO) order bias are presented on the table.

Table 4: Bias convergence in Monte Carlo simulation.

Number of obs.	AME	FO MEAI	SO MEAI	FO Bias	SO Bias	Rel. improv.
1000	1.4034	1.0060	1.2033	0.3974	0.2001	0.4965
10000	1.5300	1.0100	1.3900	0.5160	0.1400	0.7308
50000	1.5303	1.0080	1.3392	0.5222	0.1910	0.6342
100000	1.5343	1.0084	1.3500	0.5259	0.1843	0.6496
250000	1.5321	1.0082	1.3436	0.5239	0.1886	0.6401
500000	1.5338	1.0083	1.3488	0.5255	0.1850	0.6479

Note: Average marginal effect (AME), first (FO) and second order (SO) approximation of the marginal effect of the average individual (MEAI), first (FO) and second order (SO) bias and the relative improvement are presented on the table.

Table 5: Estimates and analysis of bias for the employment equations.

Variables	Est.	AME	FO MEAI	SO MEAI	FO Bias	SO Bias
Immigrants						
Constant	-1.3258	-0.3387	-0.5195	-0.3871	0.1808	0.0485
Log other income	-0.7741	-0.1977	-0.3033	-0.2260	0.1055	0.0283
Age	0.1260	0.1530	0.2347	0.1749	-0.0817	-0.0289
Age squared	-0.1615	-	-	-	-	-
Big city ($> 250,000$)	0.1115	0.0285	0.0437	0.0326	-0.1520	-0.0041
Number of children	-0.0170	-0.0044	-0.0067	-0.0050	0.0023	0.0006
Married/Cohabiting	0.3598	0.0919	0.1410	0.1051	-0.0490	-0.0132
YSM	4.7646	0.0122	1.8668	1.3914	-0.6496	-0.1742
YSM squared	-0.1058	-	-	-	-	-
Education (highest level):						
Upper-secondary	0.3657	0.0934	0.1433	0.1068	-0.0499	-0.0134
University	0.5363	0.1370	0.2101	0.1566	-0.0731	-0.0196
Arrival cohort:						
1970-1974	-0.2306	-0.0589	-0.0904	0.0314	-0.0673	0.0084
1975-1979	-0.2826	-0.0722	-0.1107	-0.0825	0.0385	0.0103
1980-1984	-0.3285	-0.0839	-0.1287	-0.0959	0.0448	0.0120
1985-1989	-0.3510	-0.0897	-0.1375	-0.1025	0.0479	0.0128
1990-1994	-0.7965	-0.2035	-0.3121	-0.2326	0.1086	0.0291
1995-2000	-0.6630	-0.1694	-0.2598	-0.1936	0.0904	0.0242
Geographical origin:						
Nordic	-0.8735	-0.2231	-0.3422	-0.2551	0.1191	0.0319
W. Europe (incl. EU)	-0.9631	-0.2461	-0.3774	-0.2813	0.1313	0.0352
USA	-1.3394	-0.3422	-0.5248	-0.3912	0.1826	0.0490
Eastern Europe	-1.3023	-0.3327	-0.5103	-0.3803	0.1776	0.0476
Middle East	-1.5686	-0.4007	-0.6146	-0.4581	0.2139	0.0573
Asia	-1.1450	-0.2925	-0.4486	-0.3344	0.1561	0.0419
Africa	-1.4546	-0.3716	-0.5699	-0.4248	0.1983	0.0532
Latin America	-1.1511	-0.2941	-0.4510	-0.3362	0.1569	0.0421
Natives						
Constant	-1.8781	-0.2753	-0.5145	-0.4719	0.2392	0.1966
Log other income	-0.8216	-0.1204	-0.2251	-0.2064	0.1046	0.0860
Age	0.1480	0.1599	0.2988	0.2741	-0.1389	-0.1142
Age squared	-0.1787	-	-	-	-	-
Big city	0.0801	0.0118	0.0220	0.0201	-0.0102	-0.0084
Number of children	0.0551	0.0080	0.0151	0.0139	-0.0070	-0.0058
Married/Cohabiting	0.3974	0.0583	0.1089	0.0999	-0.0506	-0.0416
Education (highest level):						
Upper-secondary	0.3803	0.0557	0.1042	0.0956	-0.0484	-0.0398
University	0.4964	0.0728	0.1360	0.1247	-0.0632	-0.0520

Note: Average marginal effect (AME), first (FO) and second order (SO) approximation of the marginal effect of the average individual (MEAI) and first (FO) and second order (SO) bias are presented on the table.

Table 6: Estimates and analysis of bias for the earnings equations.

Variables	Est.	AME	FO MEAI	SO MEAI	FO Bias	SO Bias
Immigrants						
Constant	11.5815	11.1524	11.0788	11.1330	0.0737	0.0195
Age	0.0290	1.3021	1.3165	1.3060	-0.0143	-0.0038
Age squared.	-0.0220	-	-	-	-	-
Big city ($> 250,000$)	-0.0541	-0.0181	-0.0119	-0.0165	-0.0062	-0.0016
Number of children	-0.0117	-0.0172	-0.0181	-0.0174	0.0009	0.0002
Married/Cohabiting	0.0217	0.1381	0.1581	0.1434	-0.0200	-0.0053
YSM	0.0075	2.2935	2.5583	2.3636	-0.2648	-0.0701
YSM squared	0.0298	-	-	-	-	-
Education (highest level):						
Upper-secondary	-0.0242	0.0941	0.1145	0.0995	-0.0203	-0.0054
University	0.1665	0.3401	0.3699	0.3479	-0.0298	-0.0079
Arrival cohort:						
1970-1974	0.0966	0.0220	0.0092	0.0186	0.0128	0.0033
1975-1979	0.1712	0.0797	0.0640	0.0756	0.0157	0.0042
1980-1984	0.2659	0.1597	0.1414	0.1548	0.0183	0.0048
1985-1989	0.3291	0.2155	0.1960	0.2103	0.0195	0.0052
1990-1994	0.4727	0.2150	0.1707	0.2032	0.0443	0.0117
1995-2000	0.6263	0.4118	0.3750	0.4021	0.0368	0.0097
Geographical origin:						
Nordic	-0.4172	-0.6998	-0.7484	-0.7127	0.0485	0.0128
W. Europe (incl. EU)	-0.3966	-0.7082	-0.7618	-0.7223	0.0535	0.0142
USA	-0.3288	-0.7622	-0.8367	-0.7819	0.0744	0.0197
Eastern Europe	-0.4382	-0.8596	-0.9320	-0.8788	0.0723	0.0191
Middle East	-0.5098	-1.0174	-1.1045	-1.0404	0.0872	0.0231
Asia	-0.4402	-0.8107	-0.8744	-0.8276	0.0636	0.0168
Africa	-0.4732	-0.9439	-1.0247	-0.9653	0.0808	0.0213
Latin America	-0.5268	-0.8993	-0.9633	-0.9162	0.0640	0.0169
Natives						
Constant	12.1808	11.3733	11.1341	11.3868	0.2392	-0.0135
Age	0.0043	1.4669	1.5893	1.4599	-0.1223	0.0069
Age squared	0.0080	-	-	-	-	-
Big city	-0.0708	-0.0363	-0.0261	-0.0369	-0.0102	0.0006
Number of children	-0.0445	-0.0208	-0.0138	-0.0212	-0.0070	0.0004
Married/Cohabiting	0.0260	0.1969	0.2475	0.1941	-0.0506	0.0029
Education (highest level):						
Upper-secondary	-0.0106	0.1529	0.2014	0.1502	-0.0484	0.0027
University	0.2361	0.4496	0.5128	0.4460	-0.0632	0.0036

Note: Average marginal effect (AME), first (FO) and second order (SO) approximation of the marginal effect of the average individual (MEAI) and first (FO) and second order (SO) bias are presented on the table.