University of Zurich

Socioeconomic Institute
Sozialökonomisches Institut

Working Paper No. 0309

Earnings Differentials between German and French Speakers in Switzerland

Alejandra Cattaneo and Rainer Winkelmann

November 2003

Socioeconomic Institute
University of Zurich

Working Paper No. 0309

Earnings Differentials between German and French Speakers in Switzerland

November 2003

Author's addresses
Alejandra Cattaneo
E-mail: alejandra.cattaneo@sts.unizh.ch

Rainer Winkelmann
E-mail: winkelmann@sts.unizh.ch

Publisher
Sozialökonomisches Institut
Bibliothek (Working Paper)
Rämistrasse 71
CH-8006 Zürich
Phone: +41-1-634 21 37
Fax: +41-1-634 49 82
URL: www.soi.unizh.ch
E-mail: soilib@soi.unizh.ch

Earnings Differentials between German and French speakers in Switzerland

Alejandra Cattaneo
University of Zurich

Rainer Winkelmann*
University of Zurich, CEPR, London and IZA, Bonn

November 2003

Abstract

The paper analyses the effect of mother tongue on labor market outcomes of Swiss residents. This type of analysis can shed light on an important policy question. Is the Swiss labor market well integrated, or can one find instead segmentation along language borders? Improving on previous research in this area, we use a nationally representative household survey, the Swiss Household Panel 1999 and 2000, and we explicitly account for self-selection of workers into language areas. Overall, we find no evidence to suggest that the Swiss labor market is not perfectly integrated or that internal migrants are positively selected

Keywords: language, labor market segmentation, Swiss household panel
JEL-Classification: J1, J31, J4, J7

*Address for correspondence: Department of Economics, University of Zurich, Zürichbergstr. 14, 8032 Zürich, 0041 1 634 2292, winkelmann@sts.unizh.ch; This research was funded by a grant from the Swiss National Science Foundation. An earlier version was presented at the Annual Meeting of the Verein für Socialpolitik, 30. Sept. - 2. Oct. 2003 in Zurich. Participants provided valuable comments.

1 Introduction

Economists have not taken great interest in the study of language so far, with the notable exception of the economic analysis of immigration, where investments in language skills are recognized as decisive for the integration process (Chiswick, 1991, Dustmann and Fabbri, 2003). However, this dismissive attitude may be about to change. One challenge is the surge of research at the intersection of psychology and economics that aims to establish a richer model of human behavior and motivation, opening up the black box of "given preferences". Cognitive processes – language being an integral part of it – move to center stage. Indeed, a school of thought in linguistics has argued for a long time that language determines thought and action (Whorf, 1956).

A second development is the increased recognition that social interactions are important for understanding human behavior and economic outcomes. Again, language is an integral aspect. A third question is linked to globalization: in a globalized world, is the optimal number of languages in the world greater than, or equal to, one? These examples are indicative of, though by no means exhausting, the type of language related questions economists might become interested in.

The topic of the current paper directly relates to the second and third points, the role of language as an integrating or disintegrating force of economic activity. What is the role of language, in relation to other factors such as jurisdiction and geography, to shape markets? What borders matter more, those of countries or those of shared language? Often these two coincide but almost as often they do not, as the examples of many English speaking countries on one hand, and multilingual countries such a Switzerland and Belgium on the other, illustrate.

We approach the issue of shared language in the context of labor markets, using the Swiss experience as a case study. Switzerland has four official languages: German, French, Italian and Romansh, although the latter is only spoken by a small minority. The three main languages correspond to three largely homogeneous linguistic regions. In such a multilingual country one might expect that language diversity generates segmented labor markets.

In this paper, we will test whether labor markets are integrated along language frontiers. Labor market integration can be defined in a number of ways. In a first definition "the law of one wage" holds across linguistic regions: Workers of similar skills receive the same wage, regardless of linguistic region. There are three classical mechanisms that each result in the law of one wage: mobility of workers, movements of capital, trade.

An alternative definition of integration is that the law of one wage holds within any given linguistic region when comparing native and nonnative workers, i.e., workers whose mother tongue matches the dominant language and those for whom this is not the case, usually because they moved across the language border. On one hand, this is a weaker form of integration: there can be wage equality within a linguistic region but wage dispersion across linguistic regions. On the other hand, it is stronger, since it focusses on labor mobility (and the resulting wages) as a precondition for integration, whereas the first type of integration does not necessarily require labor mobility.

We concentrate in our paper on this second form of integration because it relates specifically to the effect of language. In contrast, when comparing wages across linguistic regions, wage differentials are affected by non-linguistic factors as well, such as geography, institutions, industrial structure and the like. Thus, we compare in this paper wages of natives and nonnatives in a given linguistic region, and evaluate the extent to which language matters for earnings. We will answer this question using different methods, carefully addressing the issue of self-selection bias that can result if the nonnatives observed in the sample (i.e. those who decided to move into a region where their mother tongue does not match the dominant language) are not a random selection of all nonnatives.

The implications of such a study extend beyond the Swiss case as its methods and results can be applied to other multilingual countries such as Belgium or Canada, as well as common markets such as the European Union, where despite the cultural and linguistic differences among their members, there are no legal barriers for residents of member countries to move from one country to another.

The paper uses data from the Swiss Household Panel, which are well suited for this purpose as they contains information on the respondents' language skills, place of residence and income. There are momentarily no other representative survey data available in Switzerland, which include all the necessary information for such an analysis. The decennial census, for example, has the advantage of providing information on language spoken and region of residence for the entire population but it does not contain any information on income or earnings. Another survey done periodically in Switzerland is the Swiss Labor Force Survey. It includes detailed information on income but no information on language skills except for language spoken at the interview.

The paper is organized as follows. In Section 2, we summarize the previous literature on the economics of language paying special attention to the studies that analyze the difference in earnings among language groups in multilingual countries. In Section 3, we discuss the research topic we consider in this paper, which is the integration of the Swiss labor market. In Section 4, we present the data and describe briefly the modalities of the data collection. Section 5 introduces the methods we use to analyze the data, while the results obtained are presented in Section 6. Finally, Section 7 gives concluding remarks.

2 Previous Literature

Since the 1960s there have been many studies on the link between earnings and language among immigrants. More recent contributions to this literature include Chiswick (1991), Chiswick and Miller (1994, 1995, 1996, 2001, 2002), Dustmann and van Soest (2001) and Dustmann and Fabbri (2003). Even though some useful insights can be obtained from these studies, we will not discuss them in detail as our analysis will not consider immigrants.

More relevant to our present study and closer to the Swiss case is the prior research on language issues in Canada. This literature has been growing especially since the mid 1960s when the Royal Commission on Bilingualism and Biculturalism made public a study showing a big earnings gap in favor of Anglophones in Montreal. A number of following studies, among them Carliner (1981), Grenier (1987), Vaillancourt (1978) and Veltman (1979), confirmed that earnings of Francophones were lower than earnings of Anglophones and returns from learning English for Francophones were higher than returns from learning French for Anglophones. Different explanations have been given for this fact. Mostly, they refer to the concentration of power in the hands of English speakers in the English Canada and the disproportionate share of economic power of English speakers in Quebec. However the earnings disparities seemed to have declined considerably since the 70s.

For Switzerland there are two previous studies by Grin (1997, 1998), where he compares the socioeconomic status of Germanophones, Francophones and Italophones. For a report completed for the Federal Office of Statistics (1997) he uses data collected for a project called 'Linguistic competences in Switzerland'. The sample consists of 2400 respondents and contains information on mother tongue and competences in other languages making the distinction between listening, writing, speaking and reading skills and between four levels of knowledge for each category. In this report Grin analyses the relationship between earnings and language, separately for men and women, using adjusted residuals, multivariate regression and Oaxaca decomposition analysis. He finds that for German and French speaking men there are no significant differences in terms of earnings, whereas Italophones seem to have significantly lower earnings than Germanophones. This result was confirmed in Grin (1998), where he analyses whether speaking Italian as mother tongue involves a penalty. For women, mother tongue doesn't seem to have any influence on earnings, as there is no significant difference between earnings of Germanophone, Francophone and Italophone females. In contrast to Grin, we use more recent data from a nationally representative sample and we control for self-selection.

3 Does Language Matter?

In a multilingual country such as Switzerland, the various language regions could effectively produce a labor market that is segmented along language borders rather than fully integrated. There are a number of conceivable indicators of labor market integration. A first indicator would be the degree of wage equality and convergence across linguistic regions. The problem with this indicator is that wage differentials between regions are affected by non-linguistic factors as well, such as geography, institutions, industrial structure and the like.

A second indicator would be the extent of mobility that takes place between language regions. Mobility is not the main concern of this paper. Rather, we propose an alternative indicator of labor market integration, namely the absence of wage discrimination. In any language region, otherwise equally qualified workers should earn the same wage, regardless of their region of origin. In other words, the mother tongue should not matter. This is, of course, only a meaningful concept if we can assume that people living in Switzerland are truly bi or tri-lingual. There is some evidence that this is the case. Every Swiss child has to learn at least one further national language at school. In the German part of Switzerland, children learn French. In the French part, they learn German.

Under the perfect integration hypothesis, we should observe that people who work and live in a nonnative region should have the same earnings as those born and raised in that region, controlling for skills such as years of schooling and experience. Thus, we will compare the earnings of workers born and raised in one region with the earnings of workers who moved to that region from within another Swiss region. Note that this avoids wage comparisons across regions, comparisons that are made complicated by differences in industrial structure and in cost of living. There are a number of explanations, why markets might be segmented, i.e., why the language frontier may mean more than a language effect.

- Cultural differences between regions. Language origin can be related to certain cultural habits or prejudices that are rewarded differently in different regions

- Missing networks. The lack of social contacts in the host region can make it difficult for a person to find a job matching to his/her skills

- Differences in the school systems. When the quality of schooling differs from region to region, a person from a region with lower education quality receives a lower wage.

- The experience acquired in the region of origin could be less valued than experience obtained at the destination.

So, if we find evidence for segmentation, the next step would be to identify its origins.

4 Data

The following empirical analysis is based on data from the Swiss Household Panel (SHP). This is an annual survey of a nationally representative sample of more than 5000 households and all its members who are older than 14. The data collection is funded by the Swiss National Science Foundation, the University of Neuchâtel, and the Swiss Federal Office of Statistics. The interviews are held in French, German, or Italian, depending on the preference of the respondent. The data collection method is Computer Assisted Telephone Interviewing (CATI). We use waves 1999 and 2000, which include information on the respondents' language skills, canton of residence and other relevant socio-economic characteristics. The data from the SHP have been combined with data from the 1990 census, which provide some complementary information about regional and communal characteristics of the respondents' place of residence. We would have prefered to use the 2000 census for this purpose but these data were not yet available for public use at the time of writing this paper.

In order to find out if there are differences in earnings of native and nonnative speakers living in a certain part of Switzerland we need to delimit each linguistic region. To do this, we first determine the majority language spoken in each commune. In particular, we use information from the census on the percentage of people in each commune who speak a certain language as mother tongue, identified in the data as "first language for personal use". These percentages were obtained by dividing the number of people who have a particular mother tongue by the total population in the commune. People who live in communes where more than 50% of the population speak French were categorized as belonging to the French region. People who live in communes where more than 50% of the population speak German were considered as belonging to the German region. Individuals can then be either classified as natives if their mother tongue matches the majority language, or as nonnatives else. Individuals from communes with no language predominance were excluded from the sample. Because of the small amount of observations we have in the SHP for Italian speakers and the Italian canton, these were exclude from the analysis as well.

It could be argued that, for the purpose of our study, a more adequate definition of being native or nonnative should be based on the dominant language at the working place rather than at the place of residence. Unfortunately, this information is not available. All people living in a nonnative region and commuting to a native region would then be misclassified as nonnatives (and vice versa). We don't think that this is a major problem in practice. However, the following results have to be interpreted under the (possibly invalid) assumption that the occurrence of cases for which dominant language at the place of residence and work differ does not bias the analysis systematically.

The sample is restricted to Swiss born residents aged 18 to 65, whose annual full-time equivalent earnings are more than 10000 Francs. After deleting records with missing data, we obtain a sample of 5199 pooled observations. In the following we will report standard errors adjusted for pooling whenever appropiate.

——— Table 1 ———

In Table 1, we present the means of some selected variables, together with their estimated standard errors. We display results separately for natives, nonnatives and for the whole sample. There are 5075 individuals who live in a region where their mother tongue matches the dominant language and 124 whose mother tongue does not match the dominant language in their region of residence. The small number of nonnatives causes the standard errors of the variables' means to be larger. The 46% of the sample are women but the proportion of females is higher for nonnatives, 64% against 46% for natives. The average levels of education and experience are about 14 and 21 years respectively.

Mean earnings of nonnatives are about 3 log points (or 3 percent) below those of natives. Here and elsewhere, we interpret differences in mean log-earnings as *relative* differences in mean earnings (in levels). Strictly speaking, this interpretation is admissible only under the additional assumption of homoskedasticity and for small log-differentials (Winkelmann, 2001).

The point estimate suggests that working in a region where one's mother tongue is spoken is associate with higher earnings. However, a formal t-test for difference in means cannot reject the null of no difference in earnings of natives and nonnatives ($t =$ 0.74). So looking at overall averages, there appears to be no evidence for a segmented labor market. However, this result may be spurious, as natives and nonnatives may differ in other respects, and we should control for any other influences on earnings through multiple regression. This will be done in Section 6 of the paper.

5 Empirical Methods

To test whether native and nonnative speakers have the same earnings, *ceteris paribus*, we adjust the raw differential using standard regression techniques. In particular, our analysis will be based on a human capital earnings function of the form:

$$w = \alpha + x'\beta + \delta nonnative + \varepsilon \tag{1}$$

where w are logarithmic earnings, x is the vector of the individual's personal characteristics that affect earnings, and *nonnative* is a dummy variable that equals one if the individual's mother tongue matches the dominant language in his/her region of residence and zero otherwise. The coefficient of *nonnative*, δ, measures the so-called treatment effect. The "treatment" in this case is whether or not the individual moved from its region of origin and thereby becomes a nonnative. The null hypothesis to be tested is whether $\delta = 0$, i.e. whether language has no effect on earnings once we control for other relevant characteristics.

An issue that arises here is the potential endogeneity of the variable *nonnative*, that is the possibility that $E(\varepsilon|nonnative) \neq 0$. We could have a positive correlation between *nonnatives* and ε if individuals who move are positively self selected. In other words, the typical individual who decides to move, and thereby becomes a nonnative in terms of language, could have relatively higher earnings whether or not he chooses to move.

To further analyse this issue, we assume that the decision to become a nonnative obeys the following latent model:

$$nonnative^* = z'\gamma + u \tag{2}$$

$$nonnative = 1 \text{ if } nonnative^* > 0, \text{ i.e. } u > -z'\gamma$$

$$nonnative = 0 \text{ if } nonnative^* \leq 0, \text{ i.e. } u \leq -z'\gamma$$

where z collects all variables in x plus any other variables that affect the moving decision but not earnings, for example those that capture mobility cost. The following proxies for mobility costs are used: a binary indicator variable for being married, an indicator variable for being married to a person who has a different mother tongue, an interaction between this variable and an indicator for gender being female, and number of children.

We furthermore assume that x and z are exogenous, i.e. $E(\varepsilon|x, z) = 0$ and $E(u|z) = 0$. If u and ε are correlated direct estimation of the earnings equation with OLS would produce selectivity bias, since the regressor *nonnative* is correlated with the error term (the regressor depends on u, and u and ε are correlated). We will use various approaches to correct for the possibility that nonnatives might be positively self selected and to analyse the robustness of the results depending on the model specification.

5.1 Joint Normal Distribution and Maximum Likelihood Estimation

Assume that ε and u are bivariate normally distributed with mean zero, variances σ_ε^2 and 1, and correlation ρ. This is the well-known set-up discussed by Heckman (1974), who showed that the model parameters can be consistently estimated by either a two-step method or maximum likelihood. Under bivariate normality, the expected wage, conditional on being nonnative, is given by

$$E(w|x, z, nonnative = 1) = \alpha + x'\beta + \delta + \rho\lambda(z'\gamma) \qquad (3)$$

where λ is the inverse Mills ratio $\phi(z'\gamma)/\Phi(z'\gamma)$ and ϕ and Φ are the density and cumulative density function of the standard normal distribution, respectively. If ρ=0 there is no self-selection problem. However, if ρ >0 then nonnatives are positively selected, i.e., their wages will exceed those of otherwise similar natives.

To apply Heckman's two step method, we estimate in a first step a probit model to obtain a consistent estimator of γ. Then we construct the variable $\hat{\lambda}$ as follows:

$$\hat{\lambda} = \phi(z'\hat{\gamma})/\Phi(z'\hat{\gamma}) \text{ if } nonnative = 1$$

$$\hat{\lambda} = -\phi(z'\hat{\gamma})/(1 - \Phi(z'\hat{\gamma})) \text{ if } nonnative = 0$$

In a second step we estimate Equation (3) by OLS using $\hat{\lambda}$ instead of λ. Standard errors need to be adjusted.

An alternative method is to jointly estimate the earnings equation and the selection equation using full maximum likelihood (where the two-step estimator may be used as starting value). Identification is achieved both through functional form, and by virtue of excluding the mobility cost related variables (marital status, number of children) from the outcome equation. The disadvantage of these two methods is that they don't work if the errors of both equations are not normally distributed.

5.2 Alternatives to Heckman's method

There are a couple of possibilities to relax the strong assumptions underlying the previous model. A first possibility is to estimate the model without imposing that ε and u are bivariate normally distributed. Assume instead that only u is normally distributed. Then we can use the following two-step procedure: first, we estimate the probit model $P(nonnative) = \Phi(z'\gamma)$ and obtain the fitted probabilities $\Phi(z'\hat{\gamma})$. In a second step we estimate the following model by OLS.

$$w = \alpha + x'\beta + \delta\Phi(z'\hat{\gamma}) + u \tag{4}$$

Wooldridge (2000) refers to this as the "plug-in" model. As before, δ measures the treatment effect. The disadvantage of this method is that consistency of the OLS estimators relies on the probit model being the correct specification, i.e. on the normality of u.

If one wants to estimate without any distributional assumption, instrumental variable techniques can be applied. This of course requires that a good instrument for treatment is available. If more than one instrument is available, the validity of the instruments can be tested. We will use the following instruments: *number of kids*, *mixed marriage*, *female*×*mixed marriage* and *married*, where *married* is a marital status indicator variable and *mixed marriage* is a dummy variable equal to one for people married to someone who has a different mother tongue.

5.3 Classical Switching Regression Model

All methods so far are based on the outcome equation (1): this equation has a single error ε, which means that the gain (or loss) from moving to another region (and thus from becoming a nonnative) is the same, and equal to δ, for all persons. Therefore, if selectivity is present (if u and ε are correlated) it must be based on *absolute advantage*, as people who move would be better (or worse) everywhere. This may not be a very realistic characterization of the selection process, and we may instead be interested in a model where selection is based on *relative advantage*. To obtain such a model, consider a more general approach where each group has its own earnings equation as follows:

If nonnative=0

$$w_0 = \alpha_0 + x'\beta_0 + \varepsilon_0 \tag{5}$$

If nonnative=1

$$w_1 = \alpha_1 + x'\beta_1 + \varepsilon_1 \tag{6}$$

Let the decision to move be generated from the same model as in equation (2). This is the classical switching regressions model in which the returns to the variables included in x as well as the errors ε_0 and ε_1 of a person as native or as nonnative are potentially different. Different errors mean that a person could have above average earnings if she moved, but average (or below average) earnings if she decided to stay. People who choose to move may be the ones who gain most from moving.

In this case we estimate two models, one for natives and one for nonnatives. We use a two-step method following Heckman, Tobias and Vytlacil (2000). Assume that ε_0, ε_1 and u have a trivariate normal distribution. Then, the procedure is as follows: first, we estimate a probit model of the decision to move and we compute $\hat{\lambda}$. In a second step we estimate Equations (5) and (6) including the correspondent $\hat{\lambda}$ as another variable. If we find that the coefficients of $\hat{\lambda}$ are significant and different between groups then we have some evidence to think that nonnatives are self selected and that the decision to move is based on the comparative advantages.

A problem here is that we do not obtain an estimation of the treatment effect directly from the regression outcome. To obtain the treatment effect we need to do some extra calculations. At this point we should distinguish between two different parameters: the average treatment effect (ATE) and the effect of the treatment on the treated (TT). The ATE in our application is the expected gain or loss from moving for a randomly chosen individual and it is simply:

$$ATE \equiv E(w_1 - w_0)$$

where the expectations are taken with respect to ε and x. The TT is the average gain or loss from moving for those who actually moved and it is estimated as follows:

$$TT \equiv E(w_1 - w_0 | nonnative = 1)$$

We didn't need to make this distinction before as ATE equals TT in the case when $\varepsilon_1 = \varepsilon_0$.

6 Results

Two basic specifications of the earnings equation are reported in Table 2. The first one includes the variable *nonnative*, standard variables used in the human capital earnings function such as *years of schooling* and *experience*, a dummy variable *german region*, to see whether there is a general premium for living in the German part regardless of the language one speaks, a variable for working status and a variable for place of residence. The second specification, in column (2), adds two variables to control for community characteristics of the respondent's place of residence. These are the unemployment rate and the share of nonnatives who have the same mother tongue as the respondent in his/her commune of residence. The models are estimated using Ordinary Least Squares and pooling together both years but correcting the standard errors to account for pooling observations.

——— Table 2 ———

The coefficient of *nonnative* in column (1) is practically zero so there is no earnings premium for natives when we hold the other human capital variables constant. For *years of schooling* and *experience* the results are standard. For example, each additional year of schooling increases earnings by approximately 8%. The earnings of people living in the German part are estimated to be about 6% above those of people living in the French part. This effect is statistically significant at conventional levels. The regional wage disparities might be due to differences in industrial structure and economic activity between regions. The coefficients of the variables *full-time* and *urban* show that full-time workers have a 6% premium and that people who live in an urban area earn in average 3% more than their counterparts who live in a rural zone. The community variables added in column (2) are not significant. Therefore, we have no evidence that earnings of nonnatives are influenced by the proportion of people in the commune who speak the same language as mother tongue or by the unemployment rate. For that reason, we will use the first specification of the earnings equation from now on.

The sign of *nonnative*'s point estimate would indicate that the labor markets are indeed perfectly integrated. However, as we said in Section 5, we might be overestimating the coefficient of *nonnative* if individuals who move are positively self selected. For this reason we proceed to use the methods described in the previous section to correct for selectivity bias.

——— Table 3 ———

In columns (2) and (3) of Table 3 we report the results for Heckman's two step and maximum likelihood methods respectively. The full set of results is given in Table 5 in Appendix B. The coefficient of *nonnative* is now positive and insignificant in both cases. The coefficient of λ, ρ, is negative, which means that people who move are negatively selected, i.e. the "worst" people move. If an average person moved from his/her region of origin his/her earnings would be between 8 and 12 % higher than earnings of natives. However the coefficients are insignificant, so we do not have any evidence to conclude that there is a selectivity problem.

Table 3, column (4) and (5), reports the results of the Plug-in and IV methods. Again, the coefficient of *nonnative* is positive and insignificant in both cases. As mentioned in Section 5.2, we use marital status related variables as instruments for the IV part. These are standard instruments as they are thought to influence the decision to move without affecting earnings directly. Indeed, a Sargan test for overidentification cannot reject the null hypothesis that the variables are valid instruments.

When we perform a Hausman test to compare the coefficients from the OLS model with the coefficients from 2SLS we find that we cannot reject the null hypothesis of no difference between the coefficient ($\chi^2 = 0.06$). This means that as before we have no evidence to suggest that $E(\varepsilon|nonnative) \neq 0$. Therefore, confirming the results from the Heckman models, we find no evidence of self-selection.

However, these results could be due to the fact that the models are misspecified or too restrictive. Next, we consider the possibility that people who move have a comparative advantage rather than an absolute advantage, i.e. we allow the error terms ε of both groups to differ. To do so we estimate a switching regression model.

The results from the estimation of the natives and nonnatives equations are reported in Table 4, column (1) and (2) respectively, where we use the same specification of the earnings equation that we use for the OLS estimation. The full set of results is given in Table 6 in Appendix B.

——— Table 4 ———

After performing an F-test to compare the coefficients from the natives and nonnatives equations we conclude that we cannot reject the null hypothesis that the coefficients from both equations are the same. There is no evidence that β differs between the two groups. Therefore we estimate another model where we allow only the errors to be different, which simplifies the analysis as we obtain again an estimation of the treatment effect parameter directly from the regression outcome. This is reported in column (3) of Table 4.

The coefficient of *nonnative* is once again positive and insignificant. The sign of ρ_1, that is the sign of the correlation between ε_1 and u, is negative, which suggests that people who loose more from moving are the ones who move (i.e., there is selection based on comparative disadvantage). This wouldn't make sense if people's decision to move were based only on income considerations. Instead, this finding suggests that the individual decision to move is mostly based on other, not income-related, factors.

Moreover, the results reported in column (3) show that the difference between λ_0 and λ_1 is not significant so we cannot reject the null hypothesis that $\varepsilon_1 = \varepsilon_0$. Thus, we cannot reject the hypothesis that either Heckman's or the 2SLS methods are appropriate. Furthermore, as we cannot reject the null hypothesis of no selectivity we have no argument against using OLS.

To repeat, the perfect integration hypothesis says that assuming perfect bilingualism among the Swiss population, equally qualified workers should earn the same, regardless of their mother tongue. Depending on the detailed specification, the coefficient of *nonnative* is either zero (OLS) or positive but insignificant (after controlling for selectivity bias). The lack of precision is clearly affected by the small sample size, with only 124 observations on nonnatives. How strong evidence is this with regard to the perfect integration hypothesis? Note that we would require the finding of a significant negative coefficient in order to reject perfect integration. Thus, at a minimum, the evidence is compatible with integrated labor markets.

7 Conclusions

This paper was concerned with the earnings of natives and nonnatives living in the German and French linguistic regions of Switzerland. The hypothesis says that equally qualified workers residing in the same region should have the same earnings no matter their mother tongue if the labor market was perfectly integrated.

The evidence for the empirical analysis was based on data from the first and second waves of the Swiss household panel for the years 1999 and 2000. In the study we used a human capital earnings function to which we add specific variables to control for the influence of language. The results show no evidence of a negative earnings differential between nonnatives and natives even after accounting for self-selection of internal migrants. Therefore we cannot reject the hypothesis that the Swiss labor market is perfectly integrated. On the other hand, we did not find any evidence to suggest that people who move are positively self selected. Indeed, individuals who move don't seem to have neither an absolute nor a comparative advantage. This fact seems to suggest that people move influenced by factors other than earnings.

This paper is the first study of its kind using data from the Swiss Household Panel. Because we consider the topic of this paper of substantial policy concern, the question of the ways in which the individual's mother tongue can affect one's economic wealth still needs to be examined more deeply. In order to deepen the analysis, we plan to use in further studies data on regional and communal characteristics from the census 2000 instead of 1990 as well as a data set with more information about Italian speakers and the Italian canton in order to include them in the analysis.

References

Bundesamt für Statistik (2002), *Statistisches Jahrbuch der Schweiz.*

Carliner, G. (1981), "Wage Differences by Language Group and the Market for Language Skills in Canada", *The Journal of Human Resources*, 16: 384-399.

Chiswick, B. (1991), "Speaking, Reading, and Earnings among Low-Skilled Immigrants", *Journal of Labor Economics*, 9: 149-170.

Chiswick, B. and Miller, P. (1994), "Language Choice among Immigrants in a Multilingual Destination", *Journal of Population Economics*, 7: 119-131.

Chiswick, B. and Miller, P. (1995), "The Endogeneity between Language and Earnings: International Analysis", *Journal of Labor Economics*, 13: 246-288.

Chiswick, B. and Miller, P. (1996), "Ethnic Networks and Language Proficiency among Immigrants", *Journal of Population Economics*, 9: 19-35.

Chiswick, B and Miller, P. (2001), "A Model of Destination-Language Acquisition: Application to Male Immigrants in Canada", *Demography*, 38: 391-409.

Chiswick, B and Miller, P. (2002), "Immigrant earnings: Language skills, linguistic concentrations and the business cycle", *Journal of Population Economics*, 15: 31-57.

De Coulon, A. (2001), "Wage Differentials between Ethnic Groups in Switzerland", *Review of Labor Economics and Industrial Relations*, 1: 111-132.

Dustmann, C. and van Soest A. (2001) "Language and Earnings of Immigrants", *Industrial and Labor Relations Review*, 55: 473-492.

Dustmann, C. and Fabbri F.(2003) "Language Proficiency and Labor Market Performance of Immigrants in the UK", *Economic Journal*, forthcoming.

Grenier, G. (1987), "Earnings by Language Group in Quebec in 1980 and Emigration from Quebec between 1976 and 1981", *Canadian Journal of Economics*, 20: 774-791.

Grin, F. (1996), "Economic approaches to language and language planning", *International Journal of the Sociology of Language*, 121,1-203.

Grin, F. (1997), *Langue et differentials de statut socio-conomique en Suisse.* Bundesamt für Statistik, Bern.

Grin, F. and Sfreddo C. (1998), "Language-Based Earnings Differentials on the Swiss labor Market: Is Italian a Liability?", *International Journal of Manpower*, 19: 520-532.

Heckman, J. (1974), "Shadow Prices, Market Wages, and Labor Supply", *Econometrica*, 42: 679-694.

Heckman, J., Tobias, J. and Vytlacil, E. (2000), "Simple Estimators for Treatment Parameters in a Latent Variable Framework with an Application to Estimating the Returns to Schooling", *NBER Working Paper*, No. 7950.

Vaillancourt F. (1978), "Revenue et Langues, Quebec, 1961-71 ", *Journal of Canadian Studies*, 13: 63-69.

Veltman, C., Boulet, J.A. and Castonguay, C. (1979), "The Economic Context of Bilingualism and Language Transfer in the Montreal Metropolitan Area", *Canadian Journal of Economics*, 12: 468-479.

Whorf, B. L. (1956), *Language, Thought and Reality: Selected Writings of Benjamin Lee Whorf*, ed. John B. Caroll, Cambride, Mass: MIT Press.

Winkelmann, R. (2001), "Correctly Interpreting the Results from a Log-Linear Regression under Heteroskedasticty - Methods and an Application to the Relative Wages of Immigrants", *Jahrbücher für Nationalökonomie und Statistik*, 221: 418-431.

Wooldridge, J. (2002), *Econometric Analysis of Cross Section and Panel Data*, Cambride, Mass: MIT Press.

Tables

Table 1: Descriptive Statistics

	Native (1)	**Nonnative** (2)	**Whole Sample** (3)
Log earnings	11.25	11.22	11.25
	(0.01)	(0.04)	(0.01)
Years of schooling	13.67	13.69	13.67
	(0.03)	(0.20)	(0.03)
Experience	21.35	24.62	21.43
	(0.15)	(0.98)	(0.15)
Female	0.46	0.64	0.46
	(0.01)	(0.04)	(0.01)
Observations	5075	124	5199

Standard errors in parentheses

Table 2: Regression estimates of earnings by OLS

	(1)	(2)
Nonnative	0.001	0.01
	(0.046)	(0.053)
Education	0.078	0.079
	(0.004)*	(0.004)*
Experience	0.039	0.039
	(0.003)*	(0.003)*
Experience^2/100	-0.06	-0.059
	(0.006)*	(0.006)*
Female	-0.234	-0.231
	(0.018)*	(0.018)*
Year 2000	0.008	0.007
	(0.009)	(0.009)
German region	0.055	0.069
	(0.015)*	(0.020)*
Urban	0.029	0.024
	(0.014)**	(0.014)+
Full-time	0.058	0.05
	(0.023)**	(0.023)**
Unemployment rate		0.009
		(0.009)
Percentage nonnatives in the commune		-0.001
		(0.002)
Constant	9.695	9.678
	(0.072)*	(0.072)*
Observations	5201	5103
R-squared	0.28	0.28

Dependent variable: logarithmic earnings
Robust standard errors in parenthesis
+ significant at 10%; ** significant at 5%; * significant at 1%
Source: SHP, Waves 1999-2000.

Table 3: Treatment Effects correcting for selection bias

	OLS (1)	Heckman 2S (2)	Heckman ML (3)	Plug-in* (4)	IV (5)
Nonnative	0.001	0.118	0.076	0.073	0.042
	(0.046)	(0.118)	(0.098)	(0.176)	(0.146)
Lambda		-0.057	-0.037		
		(0.055)	(0.037)		
Hausman Test, χ^2					0.06
Test for OI, χ^2					8.32
R-squared	0.28			0.28	0.28

Dependent variable: logarithmic earnings
Robust standard errors in parentheses
*Standard errors are not corrected for generated regressors
Source: SHP, Waves 1999-2000
n=5199

Table 4: Treatment Effects in the extended Heckman selection model

	Nonnative (1)	Native (2)	Whole Sample (3)
Nonnative			0.179
			(0.16)
$Lambda_1$	-0.171		-0.093
	(0.083)**		(0.074)
$Lambda_0$		0.05	0.041
		(0.12)	(0.121)
Observations	124	5075	5199
R-squared	0.22	0.28	0.28

See Table 2

Appendix A: Variables

Name	Description
Log earnings	The natural logarithm of the annual gross work income (in the year prior to the interview). We use the real income to account for changes in price levels. For this purpose we deflate the nominal incomes from 1998 and 1999 using the correspondent CPI published by the Swiss Federal Office of Statistics (1998=98.0, 1999=98.8). As we include part and full-time workers we need to adjust the earnings of part-time workers in order to make them comparable with earnings of full-time workers. In order to do this we compute full-time equivalent earnings dividing reported earnings by the percentage of employment.
Nonnative	Dummy variable that equals 1 for people who live in a region where their mother tongue matches the dominant language and zero otherwise
German region	Dummy variable equal to 1 if the respondent lives in the German region
Years of schooling	It was constructed assigning numbers between 7 and 18 to the variable 'highest obtained education'
Experience	= age-education-6
Female	Dummy variable equal to 1 if the respondent is a female and 0 otherwise
Year 2000	Dummy variable equal to 1 if the observation belongs to the year 2000 and 0 otherwise
Urban	Dummy variable equal to 1 if the respondent lives in an urban area and 0 otherwise
Full-time	Dummy variable equal to 1 if the person works more than 80% and 0 otherwise
Married	Dummy variable equal to 1 if the respondent is married or has a partner and 0 otherwise
Mixed Marriage	Dummy variable equal to one for people married to a person who has another mother tongue
Number of kids	Number of kids
Percentage of nonnatives in the commune	Percentage of nonnatives in the respondent's commune who have the same mother tongue as the respondent

Appendix B: Tables

Table 5: Treatment Effects correcting for selection bias

	OLS (1)	2S (2)	ML (3)	Plug-in (4)	IV (5)
Nonnative	0.001	0.118	0.076	0.073	0.042
	(0.046)	(0.118)	(0.098)	(0.176)	(0.146)
Years of schooling	0.078	0.078	0.078	0.078	0.078
	(0.004)*	(0.003)*	(0.004)*	(0.004)*	(0.003)*
Experience	0.039	0.039	0.039	0.039	0.039
	(0.003)*	(0.002)*	(0.003)*	(0.003)*	(0.002)*
Experience^2*/100*	-0.059	-0.06	-0.06	-0.061	-0.06
	(0.006)*	(0.006)*	(0.005)*	(0.006)*	(0.005)*
Female	-0.234	-0.235	-0.235	-0.235	-0.234
	(0.018)*	(0.015)*	(0.018)*	(0.018)*	(0.015)*
Year 2000	0.008	0.008	0.008	0.008	0.008
	(0.009)	(0.012)	(0.009)	(0.009)	(0.012)
German region	0.055	0.06	0.058	0.058	0.056
	(0.015)*	(0.015)*	(0.016)*	(0.017)*	(0.015)*
Urban	0.029	0.03	0.03	0.03	0.03
	(0.014)**	(0.013)**	(0.014)**	(0.014)**	(0.013)**
Full-time	0.058	0.059	0.058	0.058	0.058
	(0.023)**	(0.017)*	(0.023)**	(0.023)**	(0.017)*
Constant	9.678	9.69	9.692	9.692	9.693
	(0.072)*	(0.054)*	(0.072)*	(0.072)*	(0.054)*
R-squared	0.28			0.28	0.28

Continuation Table 5 : Estimation of the selection equation

	OLS	Heckman 2S	Heckman ML	Plug-in	IV
Years of schooling		0.028	0.027	0.028	
		(0.02)	(0.021)	(0.02)	
Experience		0.032	0.03	0.032	
		(0.017)+	(0.023)	(0.022)	
Experience^2/100		-0.038	-0.036	-0.038	
		(0.037)	(0.048)	(0.048)	
Female		0.14	0.144	0.14	
		(0.108)	(0.134)	(0.135)	
Mixed marriage		0.763	0.787	0.763	
		(0.186)*	(0.242)*	(0.236)*	
Female mixed marriage		1.044	1.016	1.044	
		(0.248)*	(0.322)*	(0.317)*	
Married		-0.417	-0.425	-0.417	
		(0.102)*	(0.134)*	(0.132)*	
Number kids		-0.058	-0.054	-0.058	
		(0.052)	(0.059)	(0.058)	
Year 2000		0.035	0.032	0.035	
		(0.086)	(0.059)	(0.059)	
German region		-0.678	-0.677	-0.678	
		(0.086)*	(0.110)*	(0.110)*	
Urban		-0.11	-0.105	-0.11	
		(0.093)	(0.093)	(0.092)	
Fulltime		-0.104	-0.102	-0.104	
		(0.115)	(0.14)	(0.139)	
Constant		-2.196	-2.175	-2.196	
		(0.385)*	(0.416)*	(0.416)*	
Lambda		-0.057	-0.037		
		(0.055)	(0.037)		
Observations	5199	5199	5199	5199	5199

Robust standard errors in parentheses
Standard errors in column (2) are adjusted for generated regressors
Source: SHP, Waves 1999-2000
n=5199
Column (2) reports the Treatment effect estimate obtained using Heckman two step method
Column(3) reports the Treatment effect estimate obtained using Heckman full maximum likelihood
Model in column (4) is estimated plugging the fitted probabilities from a probit estimation of nonnatives in the outcome regression
Model in column (5) is estimated by 2SLS

Table 6: Treatment Effects in the extended Heckman selection model

	Nonnatives (1)	Natives (2)	Whole sample (3)
Nonnative			0.179
			(0.16)
Years of schooling	0.061	0.079	0.078
	(0.030)**	(0.004)*	(0.004)*
Experience	0.01	0.04	0.039
	(0.024)	(0.003)*	(0.003)*
Experience^2	-0.014	-0.062	-0.061
	(0.047)	(0.006)*	(0.006)*
Female	-0.365	-0.231	-0.233
	(0.101)*	(0.018)*	(0.018)*
Year 2000	-0.016	0.009	0.008
	(0.08)	(0.009)	(0.009)
German region	0.106	0.051	0.052
	(0.114)	(0.018)*	(0.018)*
Urban	0.189	0.025	0.029
	(0.094)**	(0.014)+	(0.014)**
Full-time	0.018	0.061	0.059
	(0.106)	(0.023)*	(0.023)*
Lambda$_1$	-0.171		-0.093
	(0.083)**		(0.074)
Lambda$_0$		0.05	0.041
		(0.12)	(0.121)
Constant	10.639	9.688	9.694
	(0.602)*	(0.072)*	(0.072)*
Observations	124	5075	5199
R-squared	0.22	0.28	0.28

See Table 2

Working Papers of the Socioeconomic Institute at the University of Zurich

The Working Papers of the Socioeconomic Institute can be downloaded from http://www.soi.unizh.ch/research/wp/index2.html

0313	Validity of Discrete-Choice Experiments – Evidence for Health Risk Reduction Harry Telser and Peter Zweifel, October 2003, 18 p.
0312	Parental Separation and Well-Being of Youths Rainer Winkelmann, October 2003, 20 p.
0311	Re-evaluating an Evaluation Study: The Case of the German Health Care Reform of 1997 Rainer Winkelmann, October 2003, 23 p.
0310	Downstream Investment in Oligopoly Stefan Buehler and Armin Schmutzler, September 2003, 33 p.
0309	Earning Differentials between German and French Speakers in Switzerland Alejandra Cattaneo and Rainer Winkelmann, November 2003, 27 p.
0308	Training Intensity and First Labor Market Outcomes of Apprenticeship Graduates Rob Euwals and Rainer Winkelmann, September 2003, 25 p.
0307	Co-payments for prescription drugs and the demand for doctor visits – Evidence from a natural experiment Rainer Winkelmann, September 2003, 22 p.
0306	Who Integrates? Stefan Buehler and Armin Schmutzler, August 2003, 32 p.
0305	Strategic Outsourcing Revisited Stefan Buehler and Justus Haucap, July 2003, 22 p.
0304	What does it take to sell Environmental Policy? An Empirical Analysis for Switzerland Daniel Halbheer, Sarah Niggli and Armin Schmutzler, 2003, 30 p.
0303	Mobile Number Portability Stefan Buehler and Justus Haucap, 2003, 12 p.
0302	Multiple Losses, Ex-Ante Moral Hazard, and the Non-Optimality of the Standard Insurance Contract Michael Breuer, 2003, 18 p.
0301	Lobbying against Environmental Regulation vs. Lobbying for Loopholes Andreas Polk and Armin Schmutzler, 2003, 37 p.
0214	A Product Market Theory of Worker Training Hans Gersbach and Armin Schmutzler, 2002, 34 p.
0213	Weddings with Uncertain Prospects – Mergers under Asymmetric Information Thomas Borek, Stefan Buehler and Armin Schmutzler, 2002, 37 p.
0212	Estimating Vertical Foreclosure in U.S. Gasoline Supply Zava Aydemir and Stefan Buehler, 2002, 42 p.
0211	How much Internalization of Nuclear RiskThrough Liability Insurance? Yves Schneider and Peter Zweifel, 2002, 18 p.
0210	Health Care Reform and the Number of Doctor Visits ? An Econometric Analysis Rainer Winkelmann, 2002, 32p.
0209	Infrastructure Quality in Deregulated Industries: Is there an Underinvestment Problem? Stefan Buehler, Armin Schmutzler and Men-Andri Benz, 2002, 24 p.
0208	Acquisitions versus Entry: The Evolution of Concentration Zava Aydemir and Armin Schmutzler, 2002, 35 p.

0207 Subjektive Daten in der empirischen Wirtschaftsforschung: Probleme und Perspektiven.
Rainer Winkelmann, 2002, 25 p.

0206 How Special Interests Shape Policy - A Survey
Andreas Polk, 2002, 63 p.

0205 Lobbying Activities of Multinational Firms
Andreas Polk, 2002, 32 p.

0204 Subjective Well-being and the Family
Rainer Winkelmann, 2002, 18 p.

0203 Work and health in Switzerland: Immigrants and Natives
Rainer Winkelmann, 2002, 27 p.

0202 Why do firms recruit internationally? Results from the IZA International Employer Survey 2000
Rainer Winkelmann, 2002, 25 p.

0201 Multilateral Agreement On Investments (MAI) - A Critical Assessment From An Industrial Economics Point Of View
Andreas Polk, 2002, 25 p.