ÖREBRO UNIVERSITY
SWEDISH BUSINESS SCHOOL

WORKING PAPER

1/2010

"In every rank, or great or small, 'Tis industry supports us all": Romanians and ethnic Hungarians, and their wages, in transition

Daniela Andrén

Economics

"In every rank, or great or small, 'Tis industry supports us all": Romanians and ethnic Hungarians, and their wages, in transition

Daniela Andrén

Swedish Business School at Örebro University and

School of Business, Economics and Law at the University of Gothenburg,

SE 701 82 Örebro, Sweden.

E-mail: Daniela.Andren@oru.se

January 7, 2010

Abstract

Legally binding treaties or memorandums have been used over time to regulate the issue of national borders of many European countries. As a result, relatively large groups of people have become ethnic minorities in other countries. They may conserve their ethnic identities, and therefore their children may accumulate ethnic human capital (e.g., language, culture, and religion) in addition to the general human capital of the country. Therefore, they can get access to an appropriate occupation linked by tradition or other factors to their ethnic group. This paper uses estimates from a selection model with an endogenous switch among three broad types of occupational groups to analyze the composition of the wage gap between Romanians and ethnic Hungarians in Romania before and during the transition from a planned to a market economy. The results suggest that the institutional settings of the controlled economy allowed Romanians to work in occupations that gave them the best returns, while the changes during the transition years allowed ethnic Hungarians to work in occupations that gave them the best returns.

1. Introduction

Relatively large groups of people have become ethnic minorities of other countries as a result of legally binding treaties or memorandums in the wake of wars. These groups may conserve their ethnic identities, and therefore their children may accumulate ethnic human capital (e.g., language, culture, and religion) in addition to the general human capital of the country. The institutional settings and the minority rights are different under different political regimes, which might affect not only the investment in human capital, but also the role of ethnic capital, for instance in terms of wages and the choice of occupation. This study analyzes these issues in Romania, where the Hungarian minority is a result of legally binding traties in the wake of wars, and represented for many years the largest ethnic group in Romania (about 10% of the population 1930-1990 and 6% in 2002). Not only the complexity of the process of Romanian national identity formation, influenced by the territorial exchanges between Romania and Hungary for hundred of years (up to 1920), but also the communist regime's policy of "assimilation of ethnic minorities into one socialist nation" (1965-1989) were expected to have deteriorated the culture, the traditions and other characteristics of the ethnic Hungarians. Therefore, when we exlude all other ethnic minorities (Roma, Germans,

etc.),[1] we should *not* find (big) differences between the ethnic Hungarians and Romanians with respect to several aspects related to access to education, labor force participation and wages.

In general, the institutional settings of the centrally planned economy were designed to treat all citizens equally regardless of gender and ethnicity, and many of the complaints regarding discrimination of ethnic Hungarians during the communist era applied to all people living in Romania during those years. Nevertheless, ethnic Hungarians and Germans could receive education in their native languages up to and through the university, but not in all fields. This suggests that in addition to the general human capital received by all children in school (according to the unique national curriculum), children belonging to these ethnic groups could also accumulate ethnical human capital from their (ethnic) schools and at home. These differences in patterns of investment in human capital might create (cases of) occupational "specialization" within ethnic groups,[2] and hence different wage levels. Differences in average wage rate across occupations enhance and distort the overall wage differentials among groups of people. Controlling for individual char-

---

[1] The 2002 Romanian Census reports that about 89.5% of the (21.7 million) people of Romania are ethnic Romanians, about 6.6% are ethnic Hungarians, 2.5% are Roma, 0.3% are German, 0.3% are Ukrainians, and then there are smaller numbers of Slovaks, Serbs, Croats, Bulgarians, Greeks, Armenians, Jews and others.

[2] e.g., Borjas (1992); Borjas (1995); Lehrer (2004); and Richman (2006).

acteristics and observed occupational choice is not enough to hedge this distortion (Andren & Andren (2007)). The data show that, as expected, during the communist era, the difference in wages between ethnic Hungarians and Romanians was not big: ethnic Hungarians had about 5% higher wages than Romanians. During the transition years, Romanians earned on average more than ethnic Hungarians. However, the analysis of the wage differences between and within ethnic groups, using a selection model with an endogenous switch among three broad types of occupational groups, reveals some interesting results. Romanians and ethnic Hungarians were rewarded differently under different institutional settings, and part of this is due to their selection into occupational sectors. Romanians' preferences for a given occupation changed from positive sorting during the communist regime to hierarchical sorting during transition years, while the ethnic Hungarians' preferences switched from negative sorting during the communist regime to hierarchical sorting in 1994, and to positive sorting in 2000. The *occupation* component explains more of the ethnic wage gap during the communist era than during the transition years. Additionally, even though ethnic Hungarians earned more than Romanians during the communist era, the *discrimination* component suggests that ethnic Hungarians were wage discriminated during the communist regime. During the transition years, Romanians earned on average more than ethnic Hun-

3

garians, yet the discrimination component was almost zero in 1994 and negative (-0.071) in 2000, suggesting that ethnic Hungarians were not wage discriminated during the first years of transition.

2. Previous studies

The economic literature that focuses on the effects of ethnicity on different economic outcomes is relatively new. The studies at the macro level focus mainly on the effects of ethnic diversity on economic growth,[3] while the studies at the micro level focus mainly on migration and the effects of ethnicity on various labor outcomes.[4] Ethnicity is often a permanent and static social characteristic of an individual, measured in terms of country of origin, nationality, citizenship, race, and language. Two general approaches have been taken: either it is formulated as a binary exogenous variable, or it is used as a stratification criterion for obtaining samples of natives and foreign-born, blacks and whites, etc. Therefore, ethnic discrimination and other effects related to the fact that people belong to ethnic

---

[3]See Alesina & La Ferrara (2005) for a survey of the literature on the positive and negative effects of ethnic diversity on economic policies and outcomes.

[4]See Kahanec (2007) for references to empirical evidence that minority ethnic groups earn on average less than the majority population, and this earnings differential is increasing with the minority share in the population in a given region.

groups are relatively well documented.[5] At the same time, ethnic identity and its determinants are still not properly understood.[6]

Ethnicity has been modeled in the context of Becker (1981)'s home production, using human capital acquired in the home and host country as key driving forces.[7] In this perspective, the process of human capital formation may be group specific, and much of this capital (e.g., skills related to ethnic languages and religions, as well as gender and age roles) is formed within the family and community, often fairly early in life, and affects the way in which people relate to the larger society in which the ethnic group is embedded. The differences in patterns of investment in human capital might create (cases of) occupational "specialization" within ethnic groups.[8] The lower the international transferability of human capital, the sharper the decline in occupational status and the higher the earnings disadvantage of the immigrants at the time of migration. Language proficiency has a strong and

---

[5]See Altonji & Blank (1999) for a literature review on race and gender in the labor market that covers ethnic discrimination.

[6]e.g., Chiswick (2006) and Constant & Zimmermann (2007).

[7]Economists have only recently begun to model and to empirically validate this important issue (e.g., Bodenhorn & Ruebeck (2003), Duncan & Trejo (2005), Chiswick (2006) and Constant *et al.* (2006)).

[8]According to the international transferability of human capital (the predominant theoretical framework of immigrant adjustment in the labor market) the stock of an immigrant's human capital obtained in the country of origin may not be fully transferable to the requirements of the host country's labor market (e.g., Chiswick (1978), Chiswick (1986) and Duleep & Regets (1999)).

significant effect on earnings,[9] but the value of the ethnic capital might not always contribute to labor market success of the immigrants.[10] For example, ethnic Germans from the ex-communist countries who returned to Germany during the 1990s faced the same difficulties with social and economic integration as other foreign-born did in Germany.[11] It seems that the general human capital of the host country is relatively more important than the ethnic capital. Nevertheless, it might also be that the ethnic capital has a relatively higher value in the country where the general human capital was accumulated.

In the early 1990s, all countries in Central, Eastern, and Southeastern Europe started their transition from various types of planned economies to a system that combines representative democracy with market economy. Since then, not only institutions, but also citizens (belonging to both majority and minority groups) have behaved differently regarding labor market issues previously controlled by the communist regimes. However, the literature related to the impact of different changes on majority and minority populations in these countries is novice, and focuses mainly on analyzing the effects of the transition on wages and incomes of the ethnic groups. The results suggest that there are differences across countries,

---

[9] See e.g., Chiswick & Miller (1995) and Dustmann (2002).
[10] See e.g., Bauer & Zimmermann (1999) and Zimmermann (1999).
[11] See e.g., Zimmermann (1999).

and the situation of the ethnic minorities groups was not always worse than the national majority.[12]

Previous studies on wages in Romania have analyzed the (gross hourly) wages and gender discrimination in urban and rural areas (e.g., Paternostro & Sahn (1999)), hourly wages in public and private enterprises by gender in 1994-1995 (e.g., Skoufias (2003)), and the impact of schooling on net monthly earnings from 1950 to 2000 (e.g., Andren *et al.* (2005)). Andren *et al.* (2005)'s estimates do not support their hypothesis that an improvement in the relative opportunities for more educated Hungarians and Germans workers (due to their valuable language abilities) has effectively shifted their relative labor supply functions backwards. Andren & Andren (2007) analyzed both the gender wage gap, as previous studies on Romanian data, and the wage gap between occupations in general, and separately for men and women. Their results show that the wage differences were in general much higher among workers of the same gender working in different occupations than between women and men working in the same occupational group, and women experienced a larger variation of occupational wage differentials than

---

[12]In Estonia, ethnic Russians performed relatively better in the early years of the transition (Kroncke and Smith, 1999). Non-Slovenian males in Slovenia experienced a slight decline in their earnings premium, whereas non-Estonian males in Estonia earned less than their Estonian counterparts (Orazem and Vodopivec, 2000). In Bulgaria, the mean log wage differential between ethnic Bulgarians and Turks increased almost three times from 1986 to 1997 (Giddings, 2002; 2003).

men.

3. Empirical framework

As already mentioned, the economic literature suggests that the ethnic differences in patterns of investment in human capital resulted (in some cases) in occupational specialization within ethnic groups. If there are big differences in average wage rate across occupations, the occupational specialization can therefore increase and distort the overall wage differentials across ethnic groups. Given that occupation cannot be considered an exogenous factor in the wage equation, controlling for observed occupational choice and other individual characteristics is not enough to hedge this distortion. Therefore, in order to analyze the wage differences between and within ethnic groups, we use the same selection model with an endogenous switch among three broad types of occupational groups as Andren & Andren (2007). Ideally, the occupational groups should be constructed to consider ethnic concentration, since this would allow testing of the hypothesis of an occupational specialization within ethnic groups. This task requires detailed information on occupation codes, which is not always available (as in the case of the present study). Therefore, we defined the occupational groups (or sectors) based on gender concentration, a factor that explains a relatively large part of wage gaps between

8

different groups of people. Following the analysis and the motivation from Jacobs (1995), two cut points ($c_1 = 0.33$ and $c_2 = 0.67$) of gender concentration are used to define the following three sectors: 1) male-dominated (MD) occupations (i.e., all occupations with less than 34% women; 2) gender-integrated (GI) occupations (i.e., all occupations with 34-66% women); and 3) female-dominated (FD) occupations (i.e., all occupations with more that 66% women).

The model, as defined by Equations 3.1-3.5, contains four stochastic components, U_1, U_2, U_3, and ε, which presumably are related to each other if the occupational choice is endogenous.

$$Y_1 = X\beta_1 + U_1, \tag{3.1}$$

$$Y_2 = X\beta_2 + U_2, \tag{3.2}$$

$$Y_3 = X\beta_3 + U_3, \tag{3.3}$$

$$D^* = Z\gamma + \varepsilon, \tag{3.4}$$

$$D = \begin{cases} 1 & \text{if} & \varepsilon < c_1 - Z\gamma & \text{(MD)} \\ 2 & \text{if} & c_1 - Z\gamma \leq \varepsilon \leq c_2 - Z\gamma & \text{(GI)} \\ 3 & \text{if} & \varepsilon > c_2 - Z\gamma & \text{(FD)} \end{cases} \tag{3.5}$$

9

where Y_j is the market wage for sector j ($j = 1, 2$ or 3), and β_j is the associated parameter vector, which is unique for each sector. Z is a matrix containing observed factors that determine the size of the occupational propensity score, and γ is the associated parameter vector of these factors. The dependent latent variable D^* represents the propensity to choose an MD occupation, and D is the observed counterpart of the latent variable.[13] The two unknown break points c_1 and c_2 will be estimated, and may be interpreted as intercepts since Z does not include any constant.

Within this framework, a given individual i who belongs to a given ethnic group k could be in any of the given sectors. Each sector j has its own earnings-generating function, Y_j, which depends on the observed (X) and unobserved (U) characteristics of the individual, everything else equal. The occupational choice is based on the taste or the propensity for a specific occupation, specified as a linear latent variable model (Equation 3.4). This model will be estimated separately for each ethnic group.

We assume that the stochastic components are $i.i.d.$ drawings from a multivari-

---

[13]The initial definition of D is

$$D = \begin{cases} 1 & \text{if} & D^* < c_1 & \text{(MD)} \\ 2 & \text{if} & c_1 \leq D^* \leq c_2 & \text{(GI)} \\ 3 & \text{if} & D^* > c_2 & \text{(FD)} \end{cases} .$$

ate normal distribution, i.e., $(U_1, U_2, U_3, \varepsilon) \sim N(0, \Sigma)$. In principle, one can allow for any potential correlation among the stochastic components. However, for a given individual, we only observe the actual wage and the indicated occupational choice in pairs, and not simultaneously with wages in other sectors. Therefore, the observability is partial, and we have to make inference on the population based on marginal distributions that correspond to the observed data. In particular, $Cov(U_1, \varepsilon), Cov(U_2, \varepsilon),$ and $Cov(U_3, \varepsilon)$ are allowed to be non-zero, while the covariances among the residuals from the output equations are left unspecified. The variances of the earnings equations are identified, and we normalize the variance of the selection equation to 1.

In order to form the likelihood function, we make use of the marginal bivariate normal density functions for $(U_1, \varepsilon), (U_2, \varepsilon),$ and (U_3, ε) and define the following indicator variables:

$$
\delta_1 = \begin{cases} 1 & \text{if } D = 1 \\ 0 & \text{elsewhere} \end{cases}, \delta_2 = \begin{cases} 1 & \text{if } D = 2 \\ 0 & \text{elsewhere} \end{cases}, \delta_3 = \begin{cases} 1 & \text{if } D = 3 \\ 0 & \text{elsewhere} \end{cases}. \quad (3.6)
$$

Using this information, we construct the following likelihood function:

$$L = \prod_{i=1}^{N} \left[\int_{-\infty}^{c_1 - Z_i\gamma} f(U_{1i}, \varepsilon_i) d\varepsilon_i \right]^{\delta_1} \times \left[\int_{c_1 - Z_i\gamma}^{c_2 - Z_i\gamma} f(U_{2i}, \varepsilon_i) d\varepsilon_i \right]^{\delta_2} \times \left[\int_{c_2 - Z_i\gamma}^{\infty} f(U_{3i}, \varepsilon_i) d\varepsilon_i \right]^{\delta_3}.$$

$$(3.7)$$

Even though the covariances among the earnings residuals are unidentified, we can still say something about the sorting structure with respect to the occupational gender segregation by using the signs of the covariances among the corresponding earnings residuals.[14] An interesting case is when $Cov(U_1, U_3) < 0$, which corresponds to the case when $Cov(U_1, \varepsilon)$ and $Cov(U_3, \varepsilon)$ have different signs. This is known as a *comparative advantage* or *positive sorting*, and suggests that the abilities of the individual are differently rewarded in the two sectors. As this requires that the contribution of the unobserved heterogeneity increases wages in both sectors, it is necessary that $Cov(U_1, \varepsilon) > 0$ and $Cov(U_3, \varepsilon) < 0$, which means

---

[14]These covariances provide insight into the form of sorting, according to the model developed by Roy (1951), where workers choose sectors based on income maximization, leading to a self-selection assignment of workers to productive activities (or a simple way of looking at self-selection of occupations). The original Roy model, based on the assumption that log skills are normally distributed, is shown to imply that pursuit of comparative advantage in a free market reduces earnings inequality compared to the earnings distribution that would result if workers were randomly assigned to sectors (Heckman & Honore (1990)). Heckman & Sedlacek (1985) extend the Roy model to include nonpecuniary benefits in the form of occupational preferences. With variation in satisfaction, workers no longer choose sectors simply based on income maximization. With a distribution of nonpecuniary benefits, workers set a reservation level of total benefits that they could get from a given occupation. If workers could choose only between the two occupations, they would choose the occupation that yields the highest reservation level of total benefits.

that those who are rewarded better in the MD sector will be rewarded less in the FD sector. Hence, people with a high propensity to choose an MD occupation do so because of the comparative advantage that comes with state-specific skills.

The parameters for the occupational selection equation and the domain-specific wage equations were estimated simultaneously for each ethnic group. These parameters are used to compute the components of the ethnic wage gap for the whole sample (i.e., all occupations together) and by occupational sector (i.e., MD, GI, and FD occupations), as well as the occupational wage gap for each ethnic group separately. These computations follow the formal statistical technique first introduced by Oaxaca (1973) and Blinder (1973), which builds on Becker's (1957) theory of labor discrimination.[15]

4. Data and some institutional aspects

The data used in the empirical analysis is drawn from the Romanian Integrated Household Survey (RIHS). For the socialist years, 1960-1989, we use retrospective information from the 1994 survey, and for the analyzed transition years (1994 and 2000), we use the annual household survey.[16] The number of observations

---

[15]These computations are presented in Appendix A1.

[16]Although originally designed as a panel, the data do not permit linking of individual observations across all years.

that include information about the wages and explanatory variables relevant for analysis varies across the analyzed samples: about 25,500 in the 1994 sample, 17,500 in the 2000 sample, and about 12,000 in the labor force history sample. Table B.1 in the Appendix presents the descriptive statistics for the variables used in the empirical analysis, by year and ethnicity. Tables B.2 and B.3 present descriptive statistics by occupational groups and year for Romanians and ethnic Hungarians, respectively.

The net monthly wage is computed as earnings from the primary job in the previous month minus taxes and other mandatory contributions. The wage variable refers to the previous month for the 1994 and 2000 samples and the starting wage for jobs held during the period before 1994. Our concern is wage differentials rather than overall level of real wages; hence our approach of estimating repeated cross-sections involves no deflation of the dependent variable. Nevertheless, the significant inflation during the 1990s requires some within-survey-period adjustments, for which we use the monthly dummies. Figure B.1 in the appendix shows the evolution of the monthly wages of Romanians relative to those of ethnic Hungarians during the communist regime and transition period for all occupations, and by occupational group. During the communist regime, Romanians earned on average (about 5%) less than ethnic Hungarians. The difference was a little bit

14

larger for those working in MD and GI occupations. The FD sector is the only one where ethnic Hungarians earned on average (about 2.5%) less than Romanians. However, this situation changed in 1994 and in 2000, when the monthly wages of Romanians were higher than those of ethnic Hungarians in all sectors. The largest difference (about 23%) is found in the MD sector in 2000.

The occupational wage gap within the same ethnic group has also changed over time, showing some differences between the two ethnic groups. For Romanians, the average wage difference between those working in MD occupations and those working in FD occupations decreased from about 10% during the communist period to less than 5% during the transition years (Table B.2 in the appendix), while for ethnic Hungarians it increased from about 2% to about 18% (Table B.3 in the appendix). These differences suggest that it is important to take into account the selection into the occupational sector when estimating the wage functions for Romanians and ethnic Hungarians.

Another group of variables that are important when analyzing the effect of occupational selection on the domain-specific wages were the instruments for occupational choices. Some good candidates are reported in Bako *et al.* (2006)'s case study of Hungarians in Romania, where the individuals' perceptions were integrated into a broader theoretical and policy-based analysis using expert testi-

monies, statistics, and developmental plans of the region under study. For example, the predominant opinion in their study is that most important businesses in the region are led by Romanians (due to the lack of possibility of the Hungarian community to benefit from the privatization opportunities). Another opinion is that the Hungarian minority has more opportunities in the small business sector compared to Romanians (due to extra resources coming from Hungary: interpersonal relations and opening up to investors from Hungary). Additionally, the use of native language in public institutions implies that in the municipalities where the Hungarian minority is dominant, applicants for jobs in public institutions must know Hungarian and Romanian.

Both during the communist era, and in the beginning of the transition period, wages were set according to industry-specific wage grids and hence varied only with the difficulty of the job and with worker education and experience. There was, and there is, a clear "diploma" requirement (i.e., a minimum level of education) for many occupations. However, people could choose only from a given and very limited list of jobs, sometimes restricted only to the municipality or county area.[17]

---

[17]Children began *compulsory schooling* (8 years of general education, starting in the mid 1950s) at age 6 or 7. Those who graduated from schools of general education entered *vocational schools* and apprenticeship programs, or continued their education in *secondary schools* (academic, teacher training, economic, industrial and agricultural lyceums). Vocational schools were the least competitive schools and provided training in occupations linked mosly to the big industries but also to agriculture. Both vocational schools and lyceums operated at two distinct

Therefore, we argue that *highest completed level of schooling* (linked to a diploma) is an exogenous source of variation in occupational attainment that allows us to identify the causal effect of occupation. This argument is supported by empirical evidence reported by Malamud & Pop-Eleches (2008) when analyzing the effects of the 1973 educational reform, which required two additional years of studies after the mandated 8 years of schooling for all students, including those planning to choose vocational schools. They reported that men in cohorts born immediately before and after January 1, 1959 (who could be affected by this reform), received very different types of secondary education and consequently experienced quite different occupational outcomes, although they had very similar rates of labor market participation and earnings in 1992.

However, it is well known that the level and quality of education affect wages. Therefore, in order to control for the impact of the quality of education on wages and occupational attainment, respectively, we use two different groups of educational dummies. The first group, used in the wage equations, includes three educational dummies that relate to wage level: *lower*, *medium*, and *higher*. The

---

levels. The first level of lyceum ("2 years of high school") corresponded to grades 9 and 10 of compulsory general education, while the second level ("4 years of high school") encompassed general education in grades 11 and 12. Entrance to higher education in universities, institutes, academies, and conservatories was based on competitive entrance exams, open to graduates of the second level of lyceum (i.e., all those who received a baccalaureate diploma that shows that they passed the compulsory exams).

second group, used in the selection equation, includes five dummies that relate more to selection into occupations (or the minimum educational requirement for occupations): *compulsory, vocational, high school, post-high school*, and *university.* The *lower* category in the wage equation covers the "*compulsory*" (which can be 4 or 8 years) and "*vocational*" in the selection equation, *medium* covers "*high school*" and "*post-high school*", and *higher* equals "*university*". Three dummies ("*vocational*", "*2 or 4 years high school*", and "*post high school*") are more likely to afect the sorting into occupational sectors, and are therefore used as instruments. These three educational categories are mainly the result of the 1956-57 and 1973 educational reforms.[18]

5. Results

A selection model with an endogenous switch among three broad types of occupational groups defined by their gender composition, i.e., MD, GI, and FD, was estimated separately for Romanians and ethnic Hungarians. Then we used the

---

[18]In 1956 and 1957, a unified system of primary and secondary schools, known as schools of general education, was established. The overall period of schooling was raised from 10 to 11 years, and further lengthened to 12 years in 1961 by extending the length of compulsory education from 7 to 8 years. In the end of the 1960s, the government introduced an explicit mandate to implement the 10 year system of general education. The educational reforms from 1973 and the years before did not allow students to start vocational training before completing 10 years of general education.

estimates to compute four components of the ethnic wage gap (endowments; occupational segregation; selectivity; and discrimination) and three components of the occupational wage gap (endowments; selectivity; and discrimination) within each ethnic group.

5.1. Selection into occupational groups

Table 5.1 presents the estimates of the selection equations for Romanians and ethnic Hungarians, respectively.[19] Additionally, we present the variances and some covariances of error terms of the wage and selection equations, which provide useful information regarding the sorting behavior of individuals across sectors.

---

[19]Tables B.4 in the Appendix present the estimates of domain-specific (i.e., MD, GI and FD) wage equations for ethnic Romanians and ethnic Hungarians respectively.

Table 5.1: Selection equation estimates, by ethnic group

	Ethnic Romanians			Ethnic Hungarians		
	1960-89	1994	2000	1960-89	1994	2000
c_1	-0.749***	-0.313**	-0.409**	-1.525*	-1.016*	-1.020
c_2	2.139***	2.198***	2.240***	1.664**	1.868***	1.963***
Gender (women vs. men)	0.563***	0.657***	0.675***	0.655***	0.715***	0.678***
Age	0.060	-0.125*	-0.102	0.107	0.042	0.102
$Age^2/10$	0.006	0.028***	0.023*	-0.006	-0.003	-0.009
Educational level (CG: compulsory or less)						
Vocational	-0.109**	0.220***	0.153***	0.290**	0.093	-0.394***
High-school 2/4 years[#]	0.436***	0.626***	0.568***	0.617***	0.590***	0.114
After high-school	0.460**	0.732***	0.793***	0.749**	0.964***	-0.322*
University	0.123*	0.416***	0.453***	0.200	0.248*	-0.053
Region						
R5: West	0.019	0.001	0.106***	-0.873	-0.535**	-0.281
R6: North-West	0.033	0.048*	0.068**	-0.815	-0.282	-0.284
R7: Center	0.128**	0.048*	0.118***	-0.700	-0.356	-0.222
Married	-0.025	-0.070***	-0.049*	-0.069	-0.194**	0.117
Urban	0.057	0.128***	0.010	-0.083	0.033	0.175**
Children aged< 18	-0.008	-0.020**	-0.012	-0.040	-0.025	-0.055
Intergeneration	0.062	0.027	0.099***	0.023	-0.218*	0.024
Variance-covariances						
$Var(U_1)$	0.437***	0.231***	0.256***	0.151**	0.164***	0.338**
$Var(U_2)$	0.326***	0.198***	0.210***	0.384***	0.148***	0.169***
$Var(U_3)$	0.149	0.203***	0.180***	0.091*	0.400***	0.177
$Cov(U_1, \varepsilon)$	0.606***	-0.311***	-0.340***	-0.147	-0.107	0.548***
$Cov(U_2, \varepsilon)$	0.264***	-0.234***	-0.294***	-0.350***	-0.187***	0.268***
$Cov(U_3, \varepsilon)$	-0.114	-0.310***	-0.211**	0.131	-0.590***	-0.243
n	8983	23553	15936	909	1615	1298
Likelihood	-12106.7	-28862.9	-18243.8	-1160.6	-1609.7	-1202.2

Notes: The estimate is significant at the 10% level ([*]), at the 5% level ([**]), and at the 1% level ([***]). CG stands for the comparison group. These notes hold for all tables of estimates. Dummies for 5-year plan periods and three dummies for ownership were also included.

The estimated coefficients of the occupational selection (or attainment) equation indicate that the probability of working in a given occupational group (i.e., MD, GI, and FD) differs between Romanians and ethnic Hungarians. For ex-

ample, during the communist era, Romanians with a vocational education had a lower probability of choosing FD occupations, while ethnic Hungarians with the same education had a higher probability. In the 2000 sample, having post high school education had the opposite effect, increasing the probability of choosing a FD occupation for Romanians and lowering it for ethnic Hungarians. The higher education parameter is not statistically significant for ethnic Hungarians during the communist regime and in 2000, while it is significant and positive for all years for the Romanians, suggesting that those with a university degree had a higher probability of working in the FD sector. This was also the case for the effect of living in a region with an ethnic overrepresentation (R5-R7).[20] Living in a region with an ethnic overrepresentation increased the probability of working in the FD sector for the Romanians, while except for the estimate of R5 in 1994, it did not have a statistical significant impact on the occupational choice of ethnic Hungarians. This might suggest that the communist regime's specially designed policies (such as the labor migration of large numbers of Romanians into compact areas of minority residence, and the territorial reorganization in 1968) aimed to reduce the minorities' ethnic concentration, could have resulted in an occupational spe-

---

[20]The minorities (most of them ethnic Hungarians) represent 35% of the Central development region (R7) population. Ethnic Hungarians represent more than half of the population in two counties of this region (85% in Harghita, and 74% in Covasna).

cialization of the Romanians who lived in those regions. However, the estimate for living in an urban area was significant and positive in 2000, suggesting that ethnic Hungarians living in urban areas had a higher probability of working in an FD occupation in 2000. The number of children younger than 18 had a significant (and negative) effect on occupational choice only for Romanians in 1994. This might suggest that FD occupations give more flexibility with respect to working time, in combination with more (child and/or adult) support within the family. Some of our findings support some of the results reported by Bako *et al.* (2006) when analyzing the opinions of people who live in the Central Development Region (R7) of Romania, a region with an average income higher than the country's average, where the Hungarian minority is dominant (more than 80% of the population of some municipalities).

The analysis of covariances of error terms suggests that the sorting mechanism into occupations was the same for both ethnic groups only in 1994, when all covariances of interest were negative, suggesting that workers of both ethnicities tended to perform similarly in all sectors (hierarchical sorting). Given that all covariances were negative, there was a positive selection effect for those who chose to work in MD occupations, and a negative selection effect for those who worked in FD occupations (the same is true for the comparisons MD versus GI occupations and

GI versus FD occupations). During the communist period and in 2000, the covariances of interest have different signs, which suggests that the two ethnic groups' sorting into occupational sectors was consistent with the theory of comparative advantage.[21] More exactly, a person from a given ethnic group selected the sector that paid her/him more than the average worker in the given ethnic group with the same characteristics and under the same working circumstances. However, the signs of all covariances of the two ethnic groups were always the opposite, suggesting different sorting structures. The results suggest that Romanians' preferences for a given occupation changed from positive sorting during the communist regime to hierarchical sorting during the transition years, while the ethnic Hungarians' preferences switched from negative sorting during the communist regime to hierarchical sorting in 1994, and to positive sorting in 2000. The differences between Romanians and ethnic Hungarians suggest that they were rewarded differently under different institutional settings. In other words, the big changes in the controlled economy (such as nationalization, industrialization, equal access to education, and employment for all ethnicities, etc.) allowed Romanians to work in the occupations that gave them the best returns, while the changes during the

---

[21]The estimated values for $Cov(U_3, \varepsilon)$ are not statistically significant during the communist regime and in 2000 for either Romanians or Hungarians.

transition years (especially regarding the improvement of minority rights since 1997) allowed the ethnic Hungarians to work in the occupations that gave them the best returns.

5.2. Decomposing the ethnic wage gap

In order to form the ethic wage differentials, we compute the mean differences in log wages between Romanians and ethnic Hungarians, taking into account both the individual effects that drive the occupational choice (the Mills' ratios) and the effects from the selection terms. The decomposed ethnic wage differential was computed as a transformed difference between the expected wages of Romanians and ethnic Hungarians for all occupations together, and by occupational group.

5.2.1. The overall ethnic wage gap

Table 5.2 presents the evolution of the observed ethnic wage gap and its components for all occupations together. The magnitude of the overall ethnic wage gap shows that ethnic Hungarians earned on average more than Romanians during the communist era, but less during the transition years. More exactly, the ethnic Hungarians' monthly wages were about 7% higher than those of Romanians during the communist regime, and about 10% lower during the transition years.

The observables explain about half of the observed ethnic wage gap in 1994, and about 60% in 2000. The unexplained part (about 50% in 1994 and 40% in 2000) could be related to ethnic discrimination. However, the deeper decomposition, which takes into account selection into occupational sector and the unobservables of the occupation shows totally different values for the unexplained part (which might be even more related to ethnic discrimination). Interestingly, the sign of the observed ethnic gap was always the opposite of the sign of its components (ethnic) discrimination and occupation but the same as its components endowments and selectivity.

Table 5.2: Ethnic wage gap decomposition, all occupations together

	1960-1989	1994	2000
Observed	-0.070	0.087	0.104
Endowments*	-0.105	0.047	0.062
Unexplained (I+II+III)	0.034	0.040	0.042
Discrimination (I)	0.129	-0.012	-0.071
Selectivity (II)	-0.114	0.072	0.132
Occupation (III)	0.061	-0.013	-0.017

Note: * We often refer to endowments as the component of the wage gap that is explained by the observables, i.e., the explained part of the wage gap.

Although ethnic Hungarians earned more than Romanians during the communist era, the discrimination component was positive, suggesting that ethnic Hungarians were wage discriminated during this period. During the transition

25

years, Romanians earned on average more than ethnic Hungarians, yet the discrimination component was almost zero in 1994 and negative (-0.071) in 2000, suggesting that ethnic Hungarians were not discriminated during the transition years. Due to the fact that we control for selection into occupation, we analyze the monthly wages only for individuals who worked. Even though the participation rates were almost the same for both ethnic groups during the analyzed years, we cannot exclude less or more favorable selection (due to language, migration, etc.) of the two ethnic groups in given occupations. For example, according to Bako *et al.* (2006), the requirement that applicants for jobs in public institutions must know both Hungarian and Romanian in the municipalities where the Hungarian minority is dominant, introduced during the 1990s, and percived by Romanians living in the Central Development Region (DR 7) as a discriminatory.

The *Occupation* component explains more of the ethnic wage gap during the communist era than during the transition years. Unfortunately, the relativly small number of observations does not allow disaggregation to a deeper level, which could possibly identify an occupational specialization within ethnic groups. However, the occupation component implied higher wages on average for Romanians during the communist period, but lower wages during the transition year.

The increasing wage gap during the transition years might be due to the large

migration, but also to employers' preferences, social networks and job requirements, which might be entry barriers to given occupations and/or establishments for both ethnic groups. The Barometer of the Ethnic Relations reveals that some characteristics of Romanians and ethnic Hungarians can be more or less attractive for employers regardless of their ethnicity.[22] However, the fact that the observed ethnic gap was positive while the discrimination component was negative supports the idea of a distinct relationship between the political and economic roles of the two ethnic groups. An active role of ethnic Hungarians (starting with December 1989) in the political arena seems to eliminate the ethnic wage discrimination.

5.2.2. The ethnic wage gap by occupational sector

Table 5.3 presents the ethnic wage gap and its components in all three occupational groups. In order to stress the importance of a deeper decomposition, the discrimination and selectivity components are reported both as separate components [(I) and (II)] and as a sum, the Unexplained (I+II) component, commonly used to explain the discrimination. The wage differential between Romanians and ethnic Hungarians and its components were different across sectors. The highest

---

[22]Started in 2000, the Barometer of Ethnic Relations (BER) is a database that contains data on ethnic relations in Romania, with a focus on the Romanian, Roma, and Hungarian ethnic groups.

observed difference was in the MD-occupations in 2000.

Table 5.3: Ethnic wage gap decomposition by occupational sector

	1960-1989	1994	2000
MD occupations			
Observed	-0.054	0.093	0.138
Endowments	-0.166	0.027	0.133
Unexplained (I+II)	0.112	0.067	0.005
Discrimination (I)	1.412	-0.181	-1.400
Selectivity (II)	-1.300	0.248	1.405
GI occupations			
Observed	-0.089	0.080	0.088
Endowments	-0.103	0.050	0.054
Unexplained (I+II)	0.014	0.031	0.033
Discrimination (I)	-0.022	0.055	0.079
Selectivity (II)	0.036	-0.024	-0.045
FD occupations			
Observed	-0.031	0.047	0.101
Endowments	-0.016	0.056	0.052
Unexplained (I+II)	-0.014	-0.009	0.048
Discrimination (I)	0.234	-0.449	-0.003
Selectivity (II)	-0.248	0.440	0.051

5.2.3. Decomposing the occupational wage gap

Table 5.4 presents the wage gaps between sectors (MD-FD; MD-GI; and FD-GI) and their components by ethnic group, and suggests that the observed differences were much higher within the ethnic groups than between them. The observed occupational wage gaps suggest that workers from both ethnic groups had on average relatively higher wages when working in MD occupations than when working in the FD or GI sectors. It seems that workers from both groups earned less when

28

working in the GI sector. However, the components of the decomposition are very different (in sign and magnitude), both between and within ethnic groups. The MD-FD wage gap was very low during the communist era (about 1% for Romanians, and 3% for the ethnic Hungarians), but increased during the transition years. In 2000, Romanians working in the MD sector earned 12% more than Romanians working in the FD sector. The difference was a little bit lower (8.3%) for ethnic Hungarians.

Table 5.4: Occupational wage gap decomposition, by ethnic group

		Ethnic Romanians			Ethnic Hungarians		
		1960-89	1994	2000	1960-89	1994	2000
E[Y(1)-Y(3)|X,Z]							
	Observed	0.009	0.097	0.120	0.033	0.051	0.083
	Endowments	-0.271	0.197	0.224	0.256	0.098	-0.247
	Unexplained (I+II)	0.280	-0.100	-0.104	-0.223	-0.048	0.330
	Occupation (I)	0.992	-1.018	-0.952	-0.241	-1.167	0.843
	Selectivity (II)	-0.712	0.918	0.849	0.018	1.119	-0.513
E[Y(1)-Y(2)|X,Z]							
	Observed	0.278	0.125	0.200	0.243	0.112	0.149
	Endowments	-0.036	0.016	0.067	0.181	0.006	-0.059
	Unexplained (I+II)	0.314	0.109	0.133	0.063	0.106	0.209
	Occupation (I)	1.273	-0.347	-0.398	-0.222	-0.076	1.134
	Selectivity (II)	-0.959	0.456	0.531	0.284	0.182	-0.925
E[Y(3)-Y(2)|X,Z]							
	Observed	0.269	0.028	0.079	0.210	0.061	0.066
	Endowments	0.056	-0.156	-0.075	0.063	-0.171	0.026
	Unexplained (I+II)	0.213	0.184	0.155	0.148	0.232	0.041
	Occupation (I)	0.460	0.645	0.473	-0.119	1.169	0.453
	Selectivity (II)	-0.247	-0.462	-0.318	0.267	-0.937	-0.413

However, the decomposition based on the estimates from the selection model

29

with an endogenous switch among three broad occupational sectors reveals different contributions of the differentials in average characteristics between the two occupational groups (i.e., the endowment effect) for the two ethnic groups. The selection and occupation components always have opposite signs, and their magnitudes are always much higher (in absolute value) than that of the observed occupational gap. The other two occupational wage gaps reveal the same pattern. These results might suggest that the institutional changes could result in differences related to language skills, accumulation of general human capital and ethnic human capital, and, to a certain extent, measurement errors.

5.2.4. Discussion

The components of the decomposition are different in magnitude, as expected, than the components computed using estimates from a regression that does not consider selection into occupational sectors but uses the occupational sectors or occupation dummies as explanatory variables. The endowments component is (much) higher, and the unexplained component is (much) lower. Even though we have an institutional motivation that supports our choice of instruments, we are aware that they might be only abstract representations of empirical descriptions. Therefore, we are aware of the realtively poor quality of (conventional) normal ap-

proximations with weak or irrelevant instruments. Additionally, if there are many (weak) instruments, standard estimators can be severely biased, and conventional methods of inference can be misleading (Bound *et al.* (1995)). With heterogeneous occupational effects, endogeneity creates severe problems for identification of population averages. Population average causal effects are only estimable under very strong assumptions about the effect of the instrument on the endogenous regressor (the constant treatment effect assumptions). Without such assumptions, we can only identify average effects for subpopulations that are induced by the instrument to change the value of the endogenous regressors. However, our econometric approach is based on a model that takes into consideration the mechanisms governing the choice of occupation. We analyze the relationship between occupation outcomes and the occupation choice mechanism, taking in account the unobservables in outcome and occupation choice equations.

6. Summary and conclusions

No study appears to have tried to assess the ethnic wage gap in Romania, and the extent to which it has changed during the transition years. Using the same empirical framework as Andren & Andren (2007), the present paper analyzes the ethnic and occupational wage gap by ethnicity during the communist regime, and

in the beginning of the transition from a controlled economy to a market economy. The covariances of error terms suggest that Romanians' preferences for a given occupation changed from positive sorting during the communist regime to hierarchical sorting during the transition. The ethnic Hungarians' preferences for a given occupation, which follow a negative sorting during the communist regime, changed to hierarchical sorting in 1994 and positive sorting in 2000. The differences between Romanians and ethnic Hungarians suggest that they respond differently to the institutional settings, and they were rewarded differently under different institutional settings. Romanians seem to have chosen better (their occupation) during the communist era, while ethnic Hungarians did in 2000.

The decomposition of the ethnic wage gap shows that the selectivity component and the discrimination component have the opposite signs, while their magnitude is almost comparable. This implies that when the decomposition does not take into account the selectivity component, the unexplained part (that used to be called discrimination) is very small. However, ethnic Hungarians are expected and appear historically to perform reasonably well in the Romanian labor market. The observables of the two ethnic groups are almost indistinguishable over time. However, our results reveal that, on average, the Romanians had lower (wage) returns to their endowments than the ethnic Hungarians during the communist era,

while the opposite occured during the transition years. The wage premium for the Romanians in 2000 was 10.4%; 6.2% is due to differences in average characteristics between the Romanians and ethnic Hungarians, and 4.2% is attributable to differentials in returns to endowments given negligible differences in selection effects between the two groups. The evolution of the Romanian-Hungarian wage gap in Romania suggests that occupational segregation, selectivity, and discrimination have different magnitudes and different directions during different institutional settings. The appropriate way to measure these components is, however, a matter of controversy. This paper relied on the Oaxaca-Blinder decomposition, with some modifications, to measure human capital, occupational segregation, and the unexplained (in part discriminatory) component of the wage gap, taking in account the selection into three (large) occupational groups. The results suggest that the selection into these occupations groups and discrimination are much stronger than the occupational segragation. Therefore, we can conclude that the Romanian-Hungarian wage gap is determined by a combination of pure discrimination and barriers to entry to high-paying occupations.

33

Örebro University and ESPE's 2008 conference for useful comments on earlier drafts of this paper, and Thomas Andrén for some of the SAS codes used to produce the estimates.

References

Alesina, Alberto, & La Ferrara, Eliana. 2005. Ethnic Diversity and Economic Performance. *Journal of Economic Literature*, **43**(3), 762–800.

Altonji, Joseph G., & Blank, Rebecca M. 1999. Race and Gender in the Labor Market. *Pages p3143 – 3259 of:* Ashenfelter, Orley, & Card, David (eds), *Handbook of labor economics. Volume 3C.* Amsterdam: Elsevier Science, North-Holland.

Andren, Daniela, & Andren, Thomas. 2007. *Occupational gender composition and wages in Romania. From planned equality to market inequality?* Working Paper 261. Göteborg University, Department of Economics.

Andren, Daniela, Earle, John S., & Sapatoru, Dana. 2005. The wage effects of schooling under socialism and in transition: Evidence from Romania, 1950 - 2000. *Journal of Comparative Economics*, **33**(2), 300–323.

Bako, Rozalia, Noica, Ruxandra, Stan, Razvan, & Pippidi, Mungiu Alina. 2006. *Regions, Minorities and European Integration: A Case Study on Hungarians in Romania*. Project Report. Romanian Academic Society (SAR).

Bauer, Thomas, & Zimmermann, Klaus F. 1999. *Occupational Mobility of Ethnic Migrants*. Discussion Paper 58. IZA.

Becker, Gary S. 1981. *A treatise on the family*. Harvard University Press.

Bodenhorn, Howard, & Ruebeck, Christopher S. 2003. *The Economics of Identity and the Endogeneity of Race*. Working Paper 9962. National Bureau of Economic Research.

Borjas, George J. 1992. Ethnic Capital and Intergenerational Mobility. *Quarterly Journal of Economics*, **107**(1), 123–150.

Borjas, George J. 1995. The Economic Benefits from Immigration. *Journal of Economic Perspectives*, **9**(2), 3–22.

Bound, John, Jaeger, David A., & Baker, Regina M. 1995. Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogeneous Explanatory Variable is Weak. *Journal of the American Statistical Association*, **90**(430), 443–450.

Chiswick, Barry R. 1978. The Effect of Americanization on the Earnings of Foreign-born Men. *Journal of Political Economy*, **86**(5), 897–921.

Chiswick, Barry R. 1986. Human Capital and the Labor Market Adjustment of Immigrants: Testing Alternative Hypotheses. *Pages 1–26 of:* Stark, Oded (ed), *Research in human capital and development. Volume 4. Migration, human capital and development.* London: JAI Press.

Chiswick, Barry R., & Miller, Paul W. 1995. The Endogeneity between Language and Earnings: International Analyses. *Journal of Labor Economics*, **13**(2), 246–288.

Chiswick, Carmel U. 2006. *The Economic Determinants of Ethnic Assimilation.* Discussion Paper 2212. IZA.

Constant, Amelie, & Zimmermann, Klaus F. 2007. *Measuring Ethnic Identity and Its Impact on Economic Behaviour.* CEPR Discussion Papers 6466. CEPR.

Constant, Amelie, Gataullina, Liliya, & Zimmermann, Klaus F. 2006. *Ethnosizing Immigrants.* CEPR Discussion Papers 5636. CEPR.

Duleep, Harriet Orcutt, & Regets, Mark C. 1999. Immigrants and Human-Capital Investment. *American Economic Review*, **89**(2), 186–191.

Duncan, Brian, & Trejo, Stephen J. 2005. *Ethnic Identification, Intermarriage, and Unmeasured Progress by Mexican Americans.* Working Paper 11423. National Bureau of Economic Research.

Dustmann, Christian. 2002. Speaking Fluency, Writing Fluency and Earnings of Migrants. *Pages 451–474 of:* Zimmermann, Klaus F., & Bauer, Thomas (eds), *The economics of migration. Volume 2. Assimilation of migrants.* Cheltenham, U.K. and Northampton, Mass.: Elgar.

Heckman, James J., & Honore, Bo E. 1990. The Empirical Content of the Roy Model. *Econometrica*, **58**(5), 1121–1149.

Heckman, James J., & Sedlacek, Guilherme L. 1985. Heterogeneity, Aggregation and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market. *Journal of Political Economy*, **93**(6), 1077–1125.

Jacobs, Sheila C. 1995. Changing Patterns of Sex Segregated Occupations throughout the Life-Course. *European Sociological Review*, **11**(2), 157–171.

Kahanec, Martin. 2007. *Ethnic Competition and Specialization.* Discussion Paper 3167. IZA.

Lehrer, Evelyn L. 2004. Religion as a Determinant of Economic and Demographic

Behavior in the United States. *Population and Development Review*, **30**(4), 707–726.

Malamud, Ofer, & Pop-Eleches, Cristian. 2008. *General Education vs. Vocational Training: Evidence from an Economy in Transition*. Working Paper 08.07. Harris School of Public Policy Studies, University og Chicago.

Paternostro, Stefano, & Sahn, David. 1999. *Wage Determination and Gender Discrimination in a Transition Economy: The Case of Romania*. Working Paper 2113. World Bank.

Richman, Barak D. 2006. How Community Institutions Create Economic Advantage: Jewish Diamond Merchants in New York. *Law and Social Inquiry*, **31**(2), 383–420.

Roy, A.D. 1951. Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers*, **3**, 135–146.

Skoufias, E. 2003. The Structure of Wages during the Economic Transition in Romania. *Economic Systems*, **27**, 345–366.

Zimmermann, Klaus F. 1999. *German Migration since 1989. Results and Perspectives*. Discussion Paper 50. IZA.

A. Appendix

A.1. Technical Appendix

A.1.1. Decomposing the ethnic wage gap

Since the early 1970s, the majority of the empirical literature on gender wage gap use Blinder-Oaxaca (BO) decomposition, a formal statistical technique first introduced by Oaxaca (1973) and Blinder (1973), which builds on Becker's (1957) theory of labor discrimination. It separates the portion of the gap resulting from differing characteristics of men and women from the portion that is not explained by these personal characteristics. Following this approach, we decompose both the ethnic wage gap and the occupational wage gap. In order to form the ethnic wage differentials, we compute the mean differences in log wages between Romanians and ethnic Hungarians, taking into account both the individual effects that drive the occupational choice (the Mills' ratios *that represent unobserved effects*) and the effects from the selection terms. Hence, the decomposed ethnic wage differential may be formed as a transformed difference between the expected wages of Romanians and ethnic Hungarians (for the entire group, and by occupational groups). For the all occupations together, the expected wages are the following:

$$E[Y_r|X_r, Z_r] = X_r\beta_r + E[U_r|X_r, Z_r] = X_r\beta_r + \theta_r\lambda_r$$

$$E[Y_h|X_h, Z_h] = X_h\beta_h + E[U_h|X_h, Z_h] = X_h\beta_h + \theta_h\lambda_h,$$

and therefore the difference in expected wages between Romanians and ethnic Hungarians is

$$E[Y_r|X_r, Z_r] - E[Y_h|X_h, Z_h] = (X_r\beta_r + \theta_r\lambda_r) - (X_h\beta_h + \theta_h\lambda_h), \qquad \text{(A.1)}$$

where Y_r and Y_h represent the log monthly wages of Romanians and ethnic Hungarians, respectively. X_r and X_h are the observables (endowments) of Romanians and ethnic Hungarians, and in the empirical analysis they will be represented by sample means. The vectors β_r and β_h represents the estimated parameters from the wage equation and λ_r and λ_h are the estimated Mills' ratio that accounts for the unobserved individual effects that drives the selection. θ_r and θ_h represent the effects from the selection terms, and are defined as the ratio $Cov(U_j, \varepsilon)/Var(\varepsilon)$. However, in this analysis we choose to normalize the variance of the selection equation so θ is simply equal to the covariance given in the ratio.

Equation A.1 would have been a simple wage differential if we had estimated just one equation for Romanians and one for ethnic Hungarians. However, due to the nature of our model we have three earnings equations for Romanians and three earnings equations for ethnic Hungarians, that is, one for each occupational

sector (Equations 3.1-3.3, page 9). Following Brown et al. (1980), we rewrite Equation A.1 as a weighted average in the following way:

$$E[Y_r|X_r, Z_r] - E[Y_h|X_h, Z_h] = \sum_{j=1}^{3} p_{rj}(X_{rj}\beta_{rj} + \theta_{rj}\lambda_{rj}) - \sum_{j=1}^{3} p_{hj}X_{hj}\beta_{hj} + \theta_{hj}\lambda_{hj}.$$

which can be rearranged as,

$$E[Y_r|X_r, Z_r] - E[Y_h|X_h, Z_h] =$$

$$= \underbrace{\sum_{j=1}^{3} p_{hj}(X_{rj} - X_{hj})\beta_{rj}}_{\text{Endowmnents}} - \underbrace{\sum_{j=1}^{3}(p_{rj} - p_{hj})X_{rj}\beta_{rj}}_{\text{Occupation}} +$$

$$+ \underbrace{\sum_{j=1}^{3}(p_{rj}\theta_{rj}\lambda_{rj} - p_{hj}\theta_{hj}\lambda_{hj})}_{\text{Selectivity}} + \underbrace{\sum_{j=1}^{3} p_{hj}X_{hj}(\beta_{rj} - \beta_{hj})}_{\text{Discrimination}},$$

where p_{rj} and p_{hj} represents shares or the probabilities to be in occupation j ($j = 1, 2$ or 3) for Romanians and ethnic Hungarians respectively. When the decomposition is made on the full sample, it is possible to decompose the total wage difference into four parts. The first component is related to endowments, and comes from differences in observables such as age, education and other socioeconomic factors important for the wage generation. The second component (or the occupational effect) is related to differences between Romanians and ethnic Hungarians in both the structure of occupational attainment, and their qualifications for the chosen occupation. The third effect (or the selectivity effect) is related to self selection into occupation that drives from the unobservables. Since

41

the occupational choice is made on the basis of the individuals preferences, skills or abilities related to different work tasks, this self selected choice could potentially affect the wages positively under the assumption that strong preferences and productivity has a positive association. If the mean selection effect for Romanians is stronger than for ethnic Hungarians, the total effect will be positive. However, if the sorting into different sectors is random, the corresponding effect would be zero. The last component comes from differences in return to observables between Romanians and ethnic Hungarians. Under the case of no discrimination, this component would be zero. However, a non zero effect could also come from lack of controlling for relevant variables, and is for that reason called unexplained.

A.1.2. Decomposing the ethnic wage gap by sector

The net gain of working in a given sector includes also non-pecuniary aspects of the job, and therefore occupational wage differentials may exist to compensate workers for pleasantness, safety, fringe benefits, and stability of the job. The decomposition within each occupational group j ($j = 1, 2$ or 3), can for obvious reasons not include any occupational effect other then the effect that comes from self selection, and is therefore given by

$$E[Y_{rj}|X_{rj}, Z_{rj}] - E[Y_{hj}|X_{hj}, Z_{hj}] = (X_{rj}\beta_r + \theta_{rj}\lambda_r) - (X_{hj}\beta_{hj} + \theta_{hj}\lambda_{hj})$$

which can be rearranged as

$$E[Y_{rj}|X_{rj}, Z_{rj}] - E[Y_{hj}|X_{hj}, Z_{hj}] = \underbrace{(X_{rj} - X_{hj})\beta_{rj}}_{\text{Endowments}} + \underbrace{(\theta_{rj}\lambda_{rj} - \theta_{hj}\lambda_{hj})}_{\text{Selectivity}} + \underbrace{X_{hj}(\beta_{rj} - \beta_{hj})}_{\text{Discrimination}}$$

$$(A.2)$$

This is the so called standard Blinder-Oaxaca decomposition.

A.1.3. Decomposing the occupational wage gap by ethnicity

The decomposition within each ethnic group for different occupational groups requires information about the average wage for each gender and each occupational group. For example, the expression for average wage for Romanians working in sector j is defined as following

$$E[Y_{rj}|X_{rj}, Z_{rj}, D = j] = X_{rj}\beta_{rj} + E[U_{rj}|X_{rj}, Z_{rj}, D = j]$$

$$= X_{rj}\beta_{rj} + \theta_{rj}\lambda_{rj}, j = 1, 2 \text{ or } 3$$

Using this expectation and Blinder-Oaxaca decomposition, we may define the occupational wage gap as following

$$E[Y_{rj}|X_{rj}, Z_{rj}] - E[Y_{ri}|X_{ri}, Z_{ri}] = \underbrace{(X_{rj} - X_{ri})\beta_{rj}}_{\text{Endowments}} + \underbrace{(\theta_{rj}\lambda_{rj} - \theta_{ri}\lambda_{ri})}_{\text{Selectivity}} + \underbrace{X_{ri}(\beta_{rj} - \beta_{ri})}_{\text{Occupation}}$$

$$(A.3)$$

where $i = 1, 2$ or 3; and $j \neq i$. The first component on the right hand side represents the wage difference between Romanians working in sector j and Romanians working in sector i due to observed and explained factors. The second component represents the selection factor and contains wage effects from unobserved individual characteristics that influence the wage of the individual. The third component represents the differences in return to different characteristics in different occupations, and it should be seen as an occupational factor that affects wages in different sectors since different factors are rewarded differently in different occupations.

B. Tables and Figures

Table B.1: Mean values and standard deviations (within the parentheses), by

ethnic group

	Ethnic Romanians			Ethnic Hungarians		
	1960-89	1994	2000	1960-89	1994	2000
Monthly wages	1395.3	134.7	1959.1	1466.8	121.3	1709.3
	(774.7)	(73.1)	(1162.9)	(797.8)	(62.7)	(825.1)
Women	0.469	0.414	0.452	0.427	0.433	0.481
Age	28.135	38.590	39.297	28.402	39.080	38.995
	(8.9)	(9.6)	(9.4)	(9.3)	(9.8)	(9.9)
Educational Level in wage equation						
Lower	0.709	0.538	0.501	0.728	0.600	0.542
High-school	0.233	0.341	0.359	0.222	0.310	0.369
University	0.057	0.120	0.140	0.050	0.090	0.089
Educational Level in selection equation						
Lower	0.457	0.211	0.081	0.454	0.269	0.105
Vocational	0.233	0.284	0.266	0.251	0.295	0.282
High-school 2/4 years[#]	0.233	0.341	0.431	0.222	0.310	0.441
After high-school	0.019	0.043	0.082	0.023	0.035	0.083
University	0.057	0.120	0.140	0.050	0.090	0.089
Married	0.797	0.793	0.764	0.770	0.749	0.723
Urban	0.546	0.693	0.725	0.584	0.669	0.662
State ownership	0.754	0.870	0.409	0.769	0.791	0.352
Long-term contract	0.142	0.945	0.957	0.139	0.936	0.965
Intergenerational	0.126	0.190	0.209	0.122	0.183	0.227
Household members	3.515	3.735	3.544	3.339	3.640	3.523
	(1.6)	(1.4)	(1.3)	(1.4)	(1.3)	(1.3)
Number children (younger than 18)	0.918	1.090	0.843	0.866	1.059	0.841
	(1.1)	(1.1)	(0.9)	(1.1)	(1.0)	(0.9)
Region						
R5	0.089	0.091	0.094	0.097	0.090	0.076
R6	0.116	0.119	0.136	0.266	0.376	0.358
R7	0.074	0.102	0.097	0.629	0.514	0.554
n	8983	23553	15936	909	1615	1298

Table B.2: Descriptive statistics by occupational sector, ethnic Romanians

|	Male-dominated occupations			Gender-integrated occupations			Female-dominated occupations		
	1960-89	1994	2000	1960-89	1994	2000	1960-89	1994	2000
Monthly wages	1566.2	147.8	2318.9	1356.6	131.6	1888.7	1538.7	130.4	1963.0
	(728.4)	(76.8)	(1491.4)	(781.7)	(73.2)	(1099.7)	(704.1)	(56.7)	(894.3)
Women	0.247	0.211	0.205	0.491	0.431	0.468	0.725	0.779	0.770
Age	27.742	38.843	39.526	28.228	38.491	39.190	27.753	38.887	39.909
	(8.5)	(9.6)	(8.9)	(9.0)	(9.7)	(9.6)	(8.7)	(9.3)	(8.9)
Education									
Lower	0.396	0.250	0.070	0.491	0.218	0.090	0.101	0.052	0.011
Vocational	0.323	0.364	0.367	0.226	0.282	0.272	0.103	0.090	0.025
High-school 2/4	0.215	0.292	0.399	0.210	0.315	0.400	0.654	0.727	0.789
After high-school	0.009	0.030	0.075	0.017	0.043	0.082	0.084	0.088	0.094
University	0.057	0.065	0.090	0.057	0.143	0.156	0.058	0.043	0.081
Married	0.829	0.826	0.821	0.794	0.786	0.752	0.770	0.777	0.775
Urban	0.555	0.585	0.673	0.532	0.714	0.730	0.733	0.765	0.773
State ownership	0.833	0.900	0.452	0.738	0.863	0.394	0.794	0.859	0.475
Long-term contract	0.178	0.962	0.967	0.126	0.940	0.953	0.286	0.957	0.981
Intergenerational	0.132	0.194	0.181	0.124	0.191	0.218	0.140	0.170	0.174
Household members	3.670	3.913	3.599	3.505	3.722	3.554	3.254	3.404	3.334
Children aged< 18	0.997	1.188	0.939	0.915	1.089	0.837	0.753	0.845	0.723
Region									
R5	0.093	0.098	0.091	0.087	0.088	0.093	0.103	0.099	0.119
R6	0.118	0.108	0.123	0.115	0.123	0.139	0.120	0.111	0.137
R7	0.058	0.088	0.082	0.076	0.108	0.099	0.086	0.092	0.106
n	1254	4627	2390	7264	17147	12286	465	1779	1260

Table B.3: Descriptive statistics by occupational sector, ethnic Hungarians

	Male-dominated occupations			Gender-integrated occupations			Female-dominated occupations		
	1960-89	1994	2000	1960-89	1994	2000	1960-89	1994	2000
Monthly wages	1672.2	135.7	1887.2	1440.9	118.1	1678.9	1512.1	128.5	1811.0
	(786.9)	(72.8)	(728.8)	(807.3)	(58.3)	(816.0)	(574.9)	(82.1)	(992.1)
Women	0.205	0.182	0.206	0.433	0.444	0.480	0.756	0.833	0.830
Age	29.273	40.000	39.175	28.259	38.918	38.823	29.156	39.019	40.620
	(10.0)	(9.6)	(9.7)	(9.2)	(9.9)	(10.0)	(9.5)	(8.8)	(8.8)
Education									
Lower	0.341	0.273	0.079	0.485	0.287	0.117	0.156	0.056	0.010
Vocational	0.307	0.403	0.405	0.255	0.295	0.291	0.067	0.074	0.030
High-school 2/4 y	0.250	0.255	0.413	0.195	0.282	0.412	0.644	0.759	0.780
After high-school	0.023	0.013	0.071	0.021	0.035	0.079	0.067	0.083	0.140
University	0.080	0.056	0.032	0.045	0.101	0.101	0.067	0.028	0.040
Married	0.784	0.831	0.722	0.772	0.731	0.718	0.711	0.787	0.780
Urban	0.659	0.597	0.571	0.566	0.672	0.658	0.756	0.787	0.820
State ownership	0.727	0.835	0.413	0.771	0.779	0.331	0.822	0.833	0.500
Long-term contract	0.159	0.965	0.968	0.131	0.931	0.963	0.222	0.935	0.980
Intergenerational	0.136	0.182	0.238	0.121	0.187	0.232	0.111	0.130	0.160
Household members	3.477	3.766	3.516	3.341	3.629	3.535	3.022	3.491	3.410
Children aged< 18	0.920	1.117	0.937	0.878	1.064	0.831	0.556	0.880	0.820
Region									
R5	0.114	0.139	0.079	0.095	0.078	0.072	0.089	0.130	0.110
R6	0.307	0.346	0.349	0.268	0.383	0.369	0.156	0.352	0.250
R7	0.580	0.498	0.556	0.629	0.522	0.550	0.733	0.454	0.590
n	88	231	126	776	1276	1072	45	108	100

Table B.4: Wage equations estimates, by occupational sector and ethnicity

	Ethnic Romanians			Ethnic Hungarians		
	1960-89	1994	2000	1960-89	1994	2000
MD-occupations						
Intercept	6.537 ***	4.014 ***	6.346 ***	4.094	3.991 ***	6.881 ***
Women	0.184 ***	-0.370 ***	-0.408 ***	-0.338 **	-0.177	0.141
Age	0.136	0.290 ***	0.441 ***	0.946 ***	0.215	0.630 **
Age²/100	-0.007	-0.036 **	-0.054 ***	-0.131 ***	-0.021	-0.069 *
High-school [1]	0.282 ***	-0.039	0.009	-0.083	0.059	0.260 **
University	0.434 ***	0.435 **	0.648 ***	0.375 **	0.755 ***	0.352 **
R5	-0.028	0.044	-0.068 **	0.001	0.094	-0.148
R6	-0.020	-0.013	-0.067 **	-0.385	0.117	-0.466
R7	0.161 ***	-0.105 ***	-0.088 ***	-0.268	-0.100	-0.453
Married	-0.015	0.060 ***	0.041 ***	0.093	-0.105	-0.052
Urban	0.093 ***	0.066 ***	0.086 ***	0.135	0.051	0.094
State ownership	0.067 **	0.015	0.134 ***	0.147	0.024	0.147 **
Long-term contract	0.038	0.119 ***	0.075	0.002	0.194	0.405 **
Intergenerational	-0.011	-0.080 ***	-0.118 ***	-0.001	-0.068	0.023
Household members	-0.006	0.002	-0.001	-0.028	0.024	0.029
GI-occupations						
Intercept	4.532 ***	4.350 ***	6.750 ***	5.073 ***	4.641 ***	6.988 ***
Women	-0.153 ***	-0.328 ***	-0.346 ***	-0.353 ***	-0.292 ***	-0.071 **
Age	0.101 *	0.279 ***	0.327 ***	-0.191	0.138	0.082
Age²/100	-0.012 *	-0.032 ***	-0.033 ***	0.028	-0.016	-0.002
Medium education	0.132 ***	0.004	0.053 ***	0.000	0.030	0.146 ***
Higher education	0.329 ***	0.385 ***	0.460 ***	0.287 ***	0.388 ***	0.455 ***
R5	0.108 ***	0.075 ***	-0.006	0.161	-0.028	-0.123
R6	0.095 ***	0.026 **	-0.019	0.129	-0.048	-0.152
R7	0.115 ***	-0.025 **	-0.042 ***	0.171	-0.169 **	-0.148
Married	0.037 **	0.028 ***	0.038 ***	-0.003	0.042	0.061
Urban	0.272 ***	0.089 ***	0.103 ***	0.311 ***	0.134 ***	0.089 ***
State ownership	0.485 ***	0.082 ***	0.078 ***	0.355 ***	-0.013	0.080 ***
Long-term contract	0.107 ***	0.068 ***	0.167 ***	0.095	0.161	0.247 ***
Intergenerational	-0.014	-0.080 ***	-0.114 ***	-0.037	-0.063	-0.070 *
Household members	0.013 ***	-0.006 **	-0.011 ***	0.033 *	-0.015 *	-0.019 *
FD- occupations						
Intercept	5.325 **	4.628 ***	6.911 ***	5.295 ***	7.267 ***	6.486 **
Women	-0.095	-0.296 ***	-0.275 ***	0.214	-0.328 ***	-0.147
Age	0.347 **	0.481 ***	0.518 ***	0.387	-0.603 *	0.940 *
Age²/100	-0.035 *	-0.057 ***	-0.057 ***	-0.054	0.088 **	-0.095
High-school	-0.004	-0.095 ***	-0.004	-0.061	-0.091	0.068
University	0.243 ***	0.135	0.164 ***	0.379	0.285	0.114
R5	-0.025	0.048	-0.107 ***	0.015	0.176	-0.121
R6	-0.115 **	0.036	0.022	-0.165	0.157	-0.186
R7	0.091	-0.026	-0.037	-0.002	0.105	-0.110
Married	-0.018	0.024	0.018	-0.069	-0.285 **	0.028
Urban	0.138 **	0.112 ***	0.074 ***	0.062	0.059	0.095
State ownership	0.021	0.002	0.030	0.067	0.043	0.059
Long-term contract	0.077 *	0.112 ***	0.250 ***	0.011	-0.198	-0.398
Intergenerational	-0.009	-0.042	-0.046	0.117	-0.597 ***	0.231
Household members	0.017	-0.003	-0.007	0.092 **	0.144 ***	-0.062

Notes: We also control for the 5-year plan periods or month, and agriculture and industry. [1] The comparison group for medium and higher education is lower education; and for R5, R6 and R7 is all other regions (predominantly populated by Ethnic Romanians).
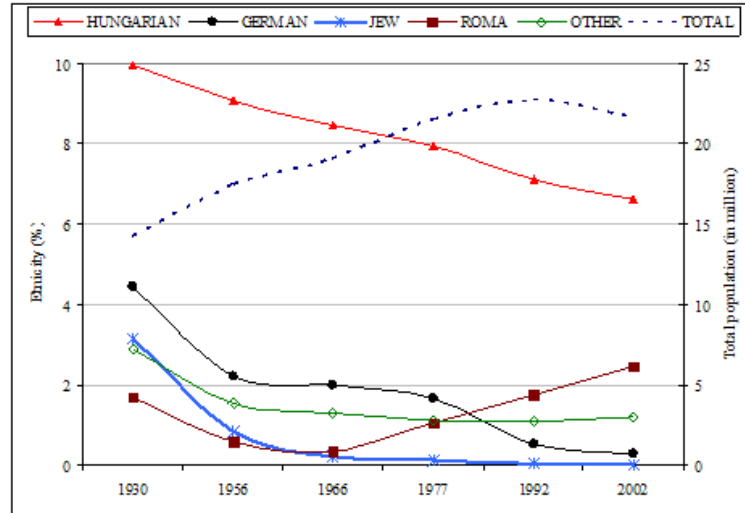
Figure B.1: Total population of Romania and the percentage of the main ethnic groups, 1930-2002
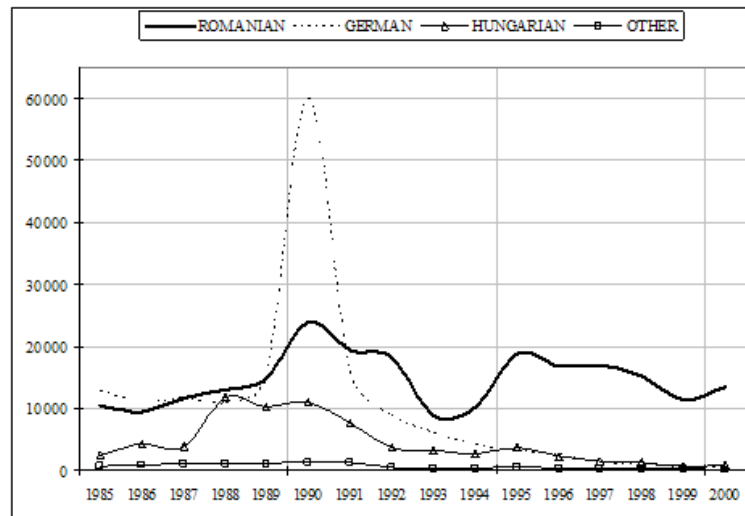
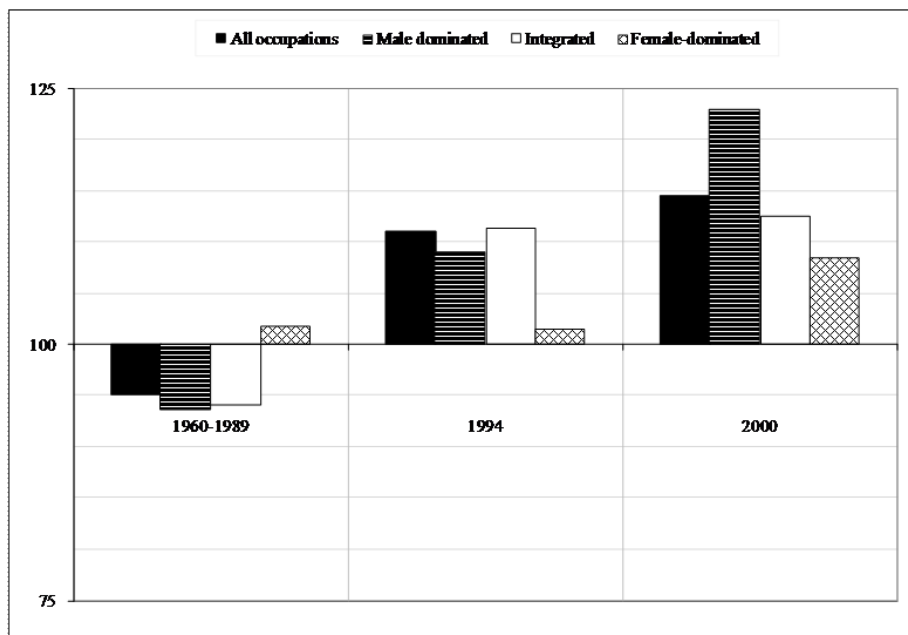Figure B.2: Emigration from Romania by ethnicity, 1985-2000

Figure B.3: The relative monthly wage of Ethnic Romanians versus Ethnic Hungarians
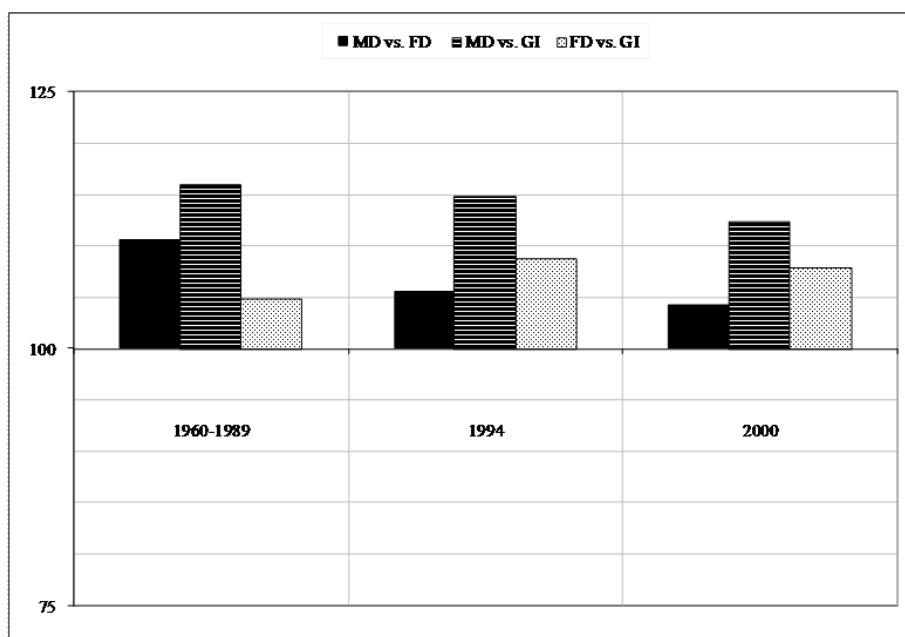
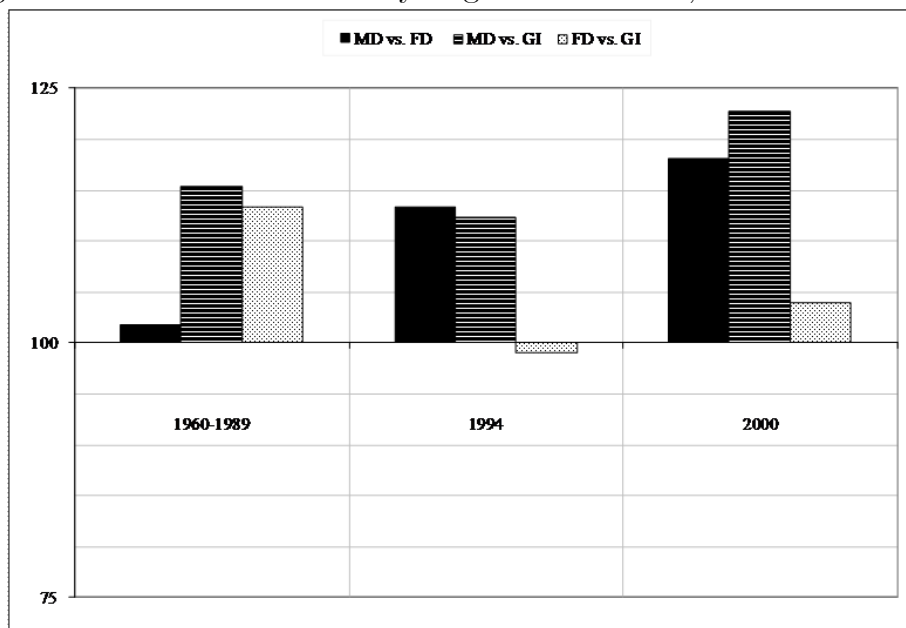Figure B.4: The relative monthly wage across sectors, ethnic Romanians

Figure B.5: The relative monthly wage across sectors, ethnic Hungarians