DISCUSSION PAPER SERIES

IZA DP No. 7667

Ethnicity and Gender in the Labour Market in Central and South East Europe

Niall O'Higgins

October 2013

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Ethnicity and Gender in the Labour Market in Central and South East Europe

Niall O'Higgins
DiSES, Università di Salerno
and IZA

Discussion Paper No. 7667
October 2013

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 7667
October 2013

ABSTRACT

Ethnicity and Gender in the Labour Market in Central and South East Europe

The Roma are both the largest 'minority' ethnic group in Central and South Eastern Europe and the one which suffered most from transition to the market. Still today, nearly forty years after the introduction of the EU's 1975 Discrimination Directive and with the end of the 'Roma Decade' (2005-15) in sight, people from the Roma minority have unemployment rates far above – and employment rates and wages far below – those of majority populations. One issue which has received relatively attention concerns the 'double' discrimination facing Roma women. Not only do Roma women face poorer employment and wage outcomes in the labour market than non-Roma women, in most CSEE countries the gender wage gap is significantly larger amongst Roma compared to non-Roma. This paper seeks to analyze and explain differences in the gender gap in the wages amongst Roma. The paper employs a non-parametric matching approach to identify the main factors underlying the gender wage gap. Educational attainment plays a relatively small role, explaining only around one-fifth of the gap. Industrial and occupational segregation appear to be playing a strong role as does the civil status of individuals, household socioeconomic status and whether individuals living in a predominantly Roma community.

JEL Classification: J16, J15

Keywords: gender, discrimination, Roma, labour market

Corresponding author:

Niall O'Higgins
DiSES
Università di Salerno
Via Ponte don Melillo
84084 Fisciano (SA)
Italy
E-mail: nohiggins@unisa.it

1. Introduction

The Roma are both the largest 'minority' ethnic group in Central and South Eastern Europe and the one which suffered most from transition to the market. Opinions differ as to the causes of these difficulties but the fact remains that still today, nearly forty years after the introduction of the EU's 1975 Discrimination Directive and with the end of the 'Roma Decade' (2005-15) in sight, people from the Roma minority have unemployment rates far above – and employment rates and wages far below – those of majority populations. Two major explanations have typically been used to account for Roma labour market disadvantage: a) the lower level of educational achievement observable amongst the Roma; and, b) the discrimination faced by Roma in the labour market. Recent work on Roma in the labour market in Central and South Eastern Europe by a number of authors has highlighted the persistence over time of ethnicity based gaps in employment and wage outcomes which cannot simply be ascribed to the albeit substantial differences in education between the Roma and non-Roma. An issue which has received less attention but which emerged in one of the more recent contributions (O'Higgins, 2012) concerns the 'double' discrimination facing Roma women in the region. Not only do Roma women face poorer employment and wage outcomes in the labour market than non-Roma women, in most CSEE countries, the gaps in labour market opportunities between Roma women and Roma men are greater than the corresponding gender gaps amongst non-Roma.

Thus, this paper focuses on intersectionality issues in the analysis of the Roma/non-Roma differences in labour market outcomes. Specifically, using data from the recent WB/UNDP regional Roma survey, the paper analyzes the gender wage gap amongst Roma and seeks to identify the main explanatory factors underlying it. In contrast to most of the preceding literature which has been based on some form of Blinder-Oaxaca decomposition, the approach adopted here is based on the non-parametric one-to-many matching approach suggested by Ñopo (2008). This innovative approach has the advantage of being both simple and at the same time particularly appropriate for analyzing Roma/non-Roma differences in labour market outcomes. The one-to-many matching

approach explicitly limits comparisons for which Roma women and men have the same characteristics and allows the separate identification of differences due to educational attainment (and other factors) and those which cannot be thus explained controlling for the problem of (the lack of) common support which would undermine the more conventional Blinder-Oaxaca type approach to gender gaps in this context.

The next section briefly reviews previous work in the area covering Roma/non-Roma labour market disadvantage and observes the complete absence of studies analysing the nature and causes of the substantial gender wage gap observable in the Roma population. The third section describes the data used for this study focusing on the sample design. Section 4 provides some descriptive evidence on the size of the gender wage gap for Roma and non-.Roma and section 5 provides a descriptive analysis of the role of education in explaining gender and ethnicity based wage gaps. Employing the aforementioned non-parametric approach, section 6 seeks to more rigorously identify the importance of education and other factors such as occupation, family characteristics and so on in explaining the gender wage gap amongst Roma, whilst final section offers some concluding remarks. Inter alia, the paper shows that the gender gap is larger for Roma than non-Roma and that, in addition to experience and educational attainment, which explain only a small portion of the gender gap, issues of occupational and industrial segregation appear to be the main drivers behind the Roma gender wage gap.

2. Previous Work

One of the difficulties with examining the factors underlying wage and employment gaps between Roma men and women is the lack of available statistical information on wages and employment subdivided by ethnicity in Central and Eastern Europe. There a number of reasons for this; inter alia National surveys typically do not – and are often not legally allowed to – subdivide the population by ethnicity. Since the new millennium systematic study of the issue has been made possible by the undertaking of several large scale studies of Roma and non-Roma. This work has

been spearheaded by the UNDP, but has also been supported by surveys undertaken by the World Bank. On the basis of these surveys, several papers have sought to identify the size – and to a lesser extent the causes – of Roma/non-Roma discrimination. Applying a standard Blinder-Oaxaca decomposition, Revenga et al. (2002, p. 21-23) found that for Roma in Bulgaria, Hungary and Romania 63% of Roma/non-Roma differences in household expenditures are explained by observed characteristics such as attained education while 37% are explained by different returns to those characteristics including discrimination. Applying Blinder–Oaxaca type decompositions (and their non-linear counterparts) to a recursive 'structural' model, O'Higgins (2010) has shown that for Roma in South-East Europe as a whole, employment and wages are less sensitive in absolute terms to educational participation resulting in a wage gain for an additional year of schooling of a little over one-third compared to the wage gain of non-Roma populations. The results also suggest that the lower returns to education for Roma – that is, the smaller benefits accruing to more educated Roma in terms of increased chances of finding work and higher wages - arising from unobserved differences attributed to discrimination, in themselves go some way towards explaining the lower educational participation of this ethnic group. Trentini (2011) applied a very similar approach to the analysis of Roma and Turk minorities in Bulgaria, finding also a proportional (and not just absolute) difference in the wage gains accruing through education between Roma and Non-Roma in that country. Also using a Blinder-Oaxaca approach adjusted for selectivity bias, Milcher (2011, p. 22) looked at the issue on a country-by-country basis and found evidence of discrimination underlying income differentials between Roma and non-Roma in Bulgaria, Kosovo and Serbia but not in Albania and Croatia. By contrast Milcher & Ficher (2011) found a significant discrimination effect explaining income differences in Albania and Kosovo but not in Bulgaria, Croatia and Serbia when applying a Bayesian approach to decomposition analysis, based upon Markov Chain Monte Carlo estimation. Applying logistic regression to Romanian census data, Kosko (2012) found that higher education significantly increases the odds of employment for Roma. Also Kosko found a "stunningly high level of unexplained difference in employment outcomes" (ibid, p. 437)

showing that Roma with the same level of education attained were more frequently unemployed and in unskilled, low wage employment.

However, all of these papers employ parametric techniques using some variation of the Blinder-Oaxaca methodology. A major problem with this type of approach is that it assumes the existence of common support, that is broadly similar characteristics, for the two groups under examination. In practice, Roma and their non-Roma neighbours differ very substantially precisely in their levels of educational attainment. Despite improvements in educational participation, Roma are still heavily concentrated in the lower ends of the educational attainment scale, whilst, non-Roma on average have much higher levels of attainment and are almost absent from the lowest 'no education' level. Given that educational attainment is also a key variable in determining wages, the assumption of similar levels of educational attainment which ia basic assumption underlying the typical parametric approach is very evidently inappropriate in this context. The matching approach adopted here, explicitly takes this issue into account.

When it comes to the issue of gender discrimination and intersectionality as it affects the double discrimination of Roma women, studies of the issue are few and far between; indeed, none to my knowledge have undertaken statistical analyses seeking to identify the extent of additional discrimination faced by Roma women or to offer empirically verified explanations for these.

3. The UNDP/WB/EC Regional Roma Survey

The statistical information used in this study are taken primarily from the UNDP/WB/EC Survey of Roma and non-Roma living in close proximity undertaken in December 2011. The survey was administered in twelve countries: Albania, Bosnia and Herzegovina, Bulgaria, Czech Republic, Croatia, Hungary, FYR of Macedonia[1], Moldova, Montenegro, Serbia, Slovakia and Romania.

[1] The Former Yugoslav Republic of Macedonia, further in the text referred to as "Macedonia".

Following the same pattern as the previous 2004 regional Roma survey conducted by UNDP, the survey contained both household and individual modules[2].

The survey was carried out via face-to-face interviews at the respondent's household, following a sampling methodology where randomly selected Roma households in areas with a high concentration of the Roma population (above the national average density of the Roma population). In parallel a control sample of non-Roma communities living in close proximity to these Roma was surveyed. In defining the Roma sample a combination of external and self-identification was used. In the analysis of the data, the term "non-Roma" relates, unless specified different, primarily to the non-Roma sample of the survey, i.e. non-Roma living in the vicinity of the surveyed Roma. The sample locations were selected from the lists of settlements, mostly from the national censuses, with average and above average shares of Roma. Although it is widely acknowledged that census data underestimate the absolute numbers of Roma, it can still be assumed that they adequately reflect the structure and territorial distribution of persons, who identify themselves as Roma.

Two/three stages random sampling was applied for both samples of the survey:

- First stage - primary sampling unit: Clusters within settlements inhabited by the Roma population (approx. size 30 households), selected by equal probability (for the Roma sample), and clusters in close proximity of settlements inhabited by the Roma population in the Roma sample (for the non-Roma sample).
- Second stage - secondary sampling unit: Households chosen with equal probabilities and selected by the method of random start and equal random walk (both samples).
- Third stage - tertiary sampling unit: Household member aged 16 and above, and selected by "first birthday" technique (both samples, only one module of the questionnaire – Module C).

The stratification was undertaken according to the type of settlements (urban/rural) and region (first sub-national level), with the purpose of optimization of the sample plan and reducing the sampling error, where the strata were defined by criteria of optimal geographical and cultural

[2] A fuller description of the survey methodology is to be found in Ivanov, A., Kling, J. and Kagin, J. (2012).

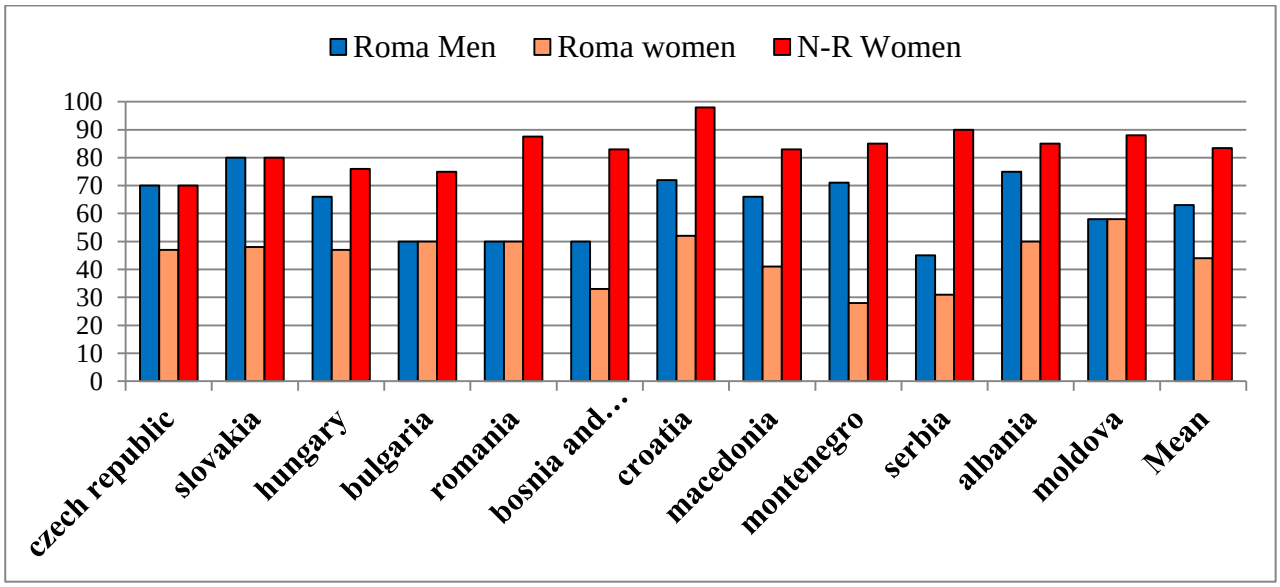
uniformity. The sample size consisted of around 750 Roma and 350 non-Roma households in each country producing a sample of 54660 individuals with each country accounting for around 4000-4500 persons.

Thus, the sample is by no means nationally representative as regards the non-Roma populations. The purpose of the survey design was indeed not intended to be so. Rather this non-representative design was intended to reduce the degree to which the two populations differed due to arbitrarily varying characteristics which would affect outcomes – the vicinity of specific industries for example or other harder to identify factors which might affect outcomes but which might not be picked up by researchers. Thus, the idea was to build into the sample design some controls for unobserved factors – particularly those related to geography - which might otherwise erroneously be attributed to ethnicity. The choice of design thus allows for sound comparisons to be made between Roma and non-Roma living in otherwise similar circumstances using a relatively small control sample of non-Roma.

4. Roma and gender in the Labour market

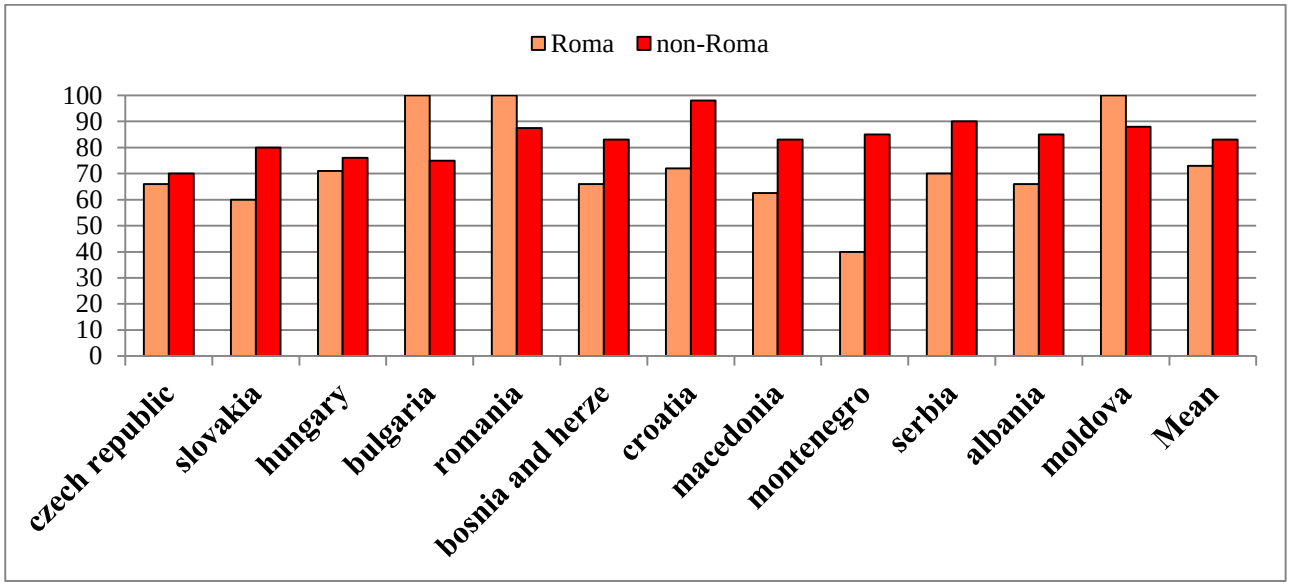
Roma men and women face substantial gaps in labour market outcomes in the countries of Central and South Eastern Europe. Whilst these disadvantages make themselves felt both of terms of (un)employment and wages, the focus in this article is on the latter. Figure 1 reports the ratio of male and female Roma and female non-Roma median monthly wages as a percentage of male non-Roma monthly pay. The comparison is limited to those in employment. The figure illustrates that the double disadvantage faced by Roma women who, with the exceptions of Bulgaria, Romania and Moldova face wages which are below those of Roma men and are everywhere well below those of non-Roma men. Indeed, the median wages of Roma women never reach even 60% of non-Roma men and in Montenegro and Serbia they are less than one-third.

Figure 1: median monthly wages of Roma men and women and non-Roma women as a percentage of the male non-Roma median monthly wage



Source: calculated on the basis of the UNDP/WB/EC regional Roma survey 2011.
Note: The figure reports the ratio of median monthly wages of male and female Roma and female non-Roma employees to non-Roma men.

Figure 2: Median monthly wages of females as a percentage of ethnicity specific male wages



Source: calculated on the basis of the UNDP/WB/EC regional Roma survey 2011.
Note: The figure reports the ratio of median monthly wages for male and female Roma employees to their non-Roma (gender-specific) counterparts.

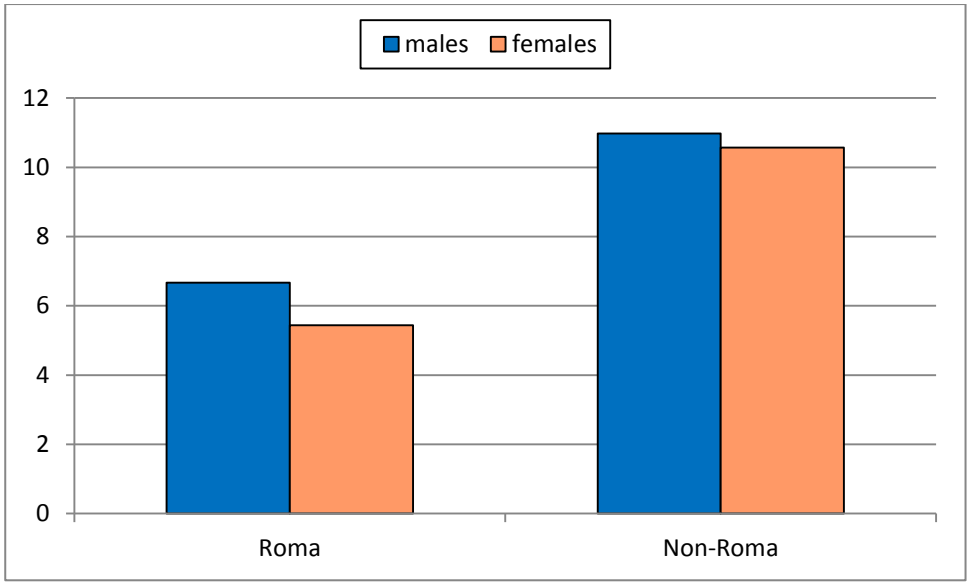
If one directly compares male and female wages by ethnicity (figure 2) one can see also that – with the already cited exceptions of Bulgaria, Romania and Moldova – Roma women have a

gender based wage gap which is significantly larger than that facing non-Roma women. The rest of this paper is primarily concerned with investigating why.

5. Too little education?

Explanations of Roma disadvantage in the labour market typically focus on one of two polarizing albeit not mutually exclusive explanations: a) the low level of educational attainment of Roma; b) the existence of labour market discrimination. Certainly educational attainment amongst the Roma is significantly lower than amongst non-Roma populations (figure 3); moreover, one can

Figure 3: Mean years of schooling by gender and ethnicity.



Source: calculated on the basis of the UNDP/WB/EC regional Roma survey 2011.

observe that the duration of educational studies is significantly shorter for Roma women than Roma men and the gender difference is greater for Roma than non-Roma. A little more detail is provided by table 1 which reports the distribution of adult Roma and non-Roma by educational attainment level. Both Roma men and women are heavily concentrated in the lower level of educational

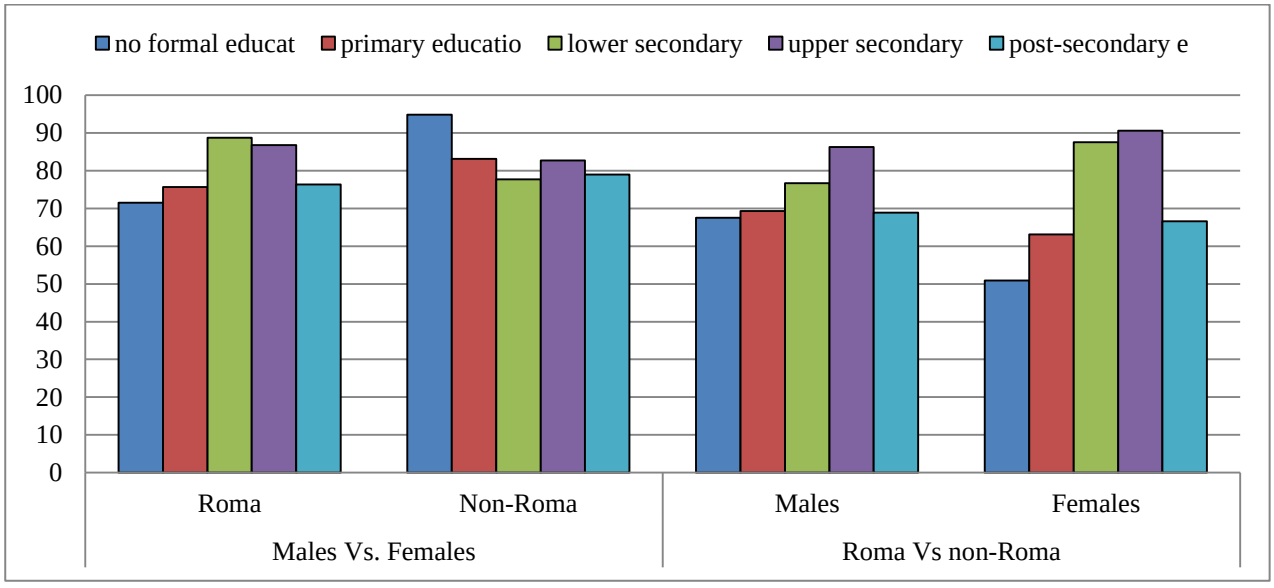
attainment and almost none have obtained post-secondary qualifications. For the non-Roma, the situation is reversed with around 1-2% of the non-Roma adult population being without any qualification, whilst 18% of Roma men and 30% of Roma women have no formal education. Perhaps these differences can provide at least a partial explanation of the additional disadvantages faced by Roma women in the labour market?

Table 1: Distribution of the sample by educational attainment by gender and ethnicity

	Roma		Non-Roma	
Highest educational level attained:	Male	Female	Male	Female
no formal education	18.3	30.6	1.4	2.3
primary education	26.2	27.0	5.0	7.5
lower secondary education	39.5	33.3	24.1	29.0
upper secondary education	15.7	8.9	61.4	51.5
post-secondary education	0.3	0.3	8.2	9.7
Total	**100.0**	**100.0**	**100.0**	**100.0**

Source: calculated on the basis of the UNDP/WB/EC regional Roma survey 2011.

Figure 4: Gender and ethnic wage gaps by education



Source: calculated on the basis of the UNDP/WB/EC regional Roma survey 2011.

Note: The figure reports the ratio of median monthly wages by education for Roma and non-Roma female employees to their (ethnicity specific) male counterparts (on the left hand side) and for male and female Roma employees to their (gender specific) non-Roma counterparts .

Examination of gender and ethnicity based differences in median wages by education provides a simple visual examination of this question (figure 4). The overall impression from the figure is that an education based explanation of differentials may be plausible for the ethnic gap; the gap narrows as educational attainment increases for both men and women – until post-secondary level is reached at which point it drops sharply. On the other hand, it does not appear to explain the gender gap – particularly for non-Roma women.

6. Towards explaining the Roma gender wage gap

In order to look at this issue a bit more carefully, I adopt here a simple non-parametric framework to examine the issue. In the past the predominant approach has been econometric estimation of some form of Mincerian returns to education equation, based on the approach suggested independently by Blinder (1973) and Oaxaca (1973) estimated separately for Roma and non-Roma (or indeed women and men). In this framework, (log) wages are regressed on indicators of educational attainment and labour market experience. A distinction is then made between differences in wages due to differences in individuals' characteristics and differences which are due to different 'returns' to specific characteristics. The first type of difference is thus explained by differences across individuals – in this case primarily differences in educational attainment levels – which the second type cannot thus be thus attributed and remains unexplained. Typically, the latter difference is then attributed to labour market discrimination.

A digression is necessary on terminology and in particular on the use of the term 'discrimination'. The usual approach is to seek to split the wage difference (or difference in some other labour market outcome) into explained and unexplained components. The idea is that the part of difference which is due to differences in characteristics -. The explained portion - is somehow justified and/or not an issue of concern, particularly if it is due to differences in educational

attainment and (potential) labour market experience. On the other hand, the unexplained portion is attributed to labour market discrimination. Such an approach is clearly overly simplistic. O'Higgins and Brüggemann (2013) have shown that the segregation of Roma into special schools for the disabled in the Czech republic and Slovakia affects Roma labour market outcomes both directly, by reducing the quality of education and hence employment opportunities and wages, but also indirectly, by impeding educational advancement. Both of these effects might reasonably be attributed to discriminatory practices, however, a typical approach would attribute the direct effect to labour market discrimination even though it may well simply be an issue of school quality, whilst remaining silent on the key issue of educational segregation which is the primary cause of ethnically based differences in outcomes – the real problem in that context. In what follows I seek to be neutral in analyzing the issue of the Roma gender wage gap in as much as I seek to identify where the problems lie – and hence where solutions may be sought - rather than attributing value laden terms to the outcomes.

Leaving aside issues of terminology, there remains the major measurement issue with using a parametric approach in this context. As noted above, econometric estimates of the Blinder-Oaxaca type assume the existence of common support, that is to say, broadly similar characteristics, for the two groups under examination. In fact, Roma and their non-Roma neighbours differ markedly in their levels of educational attainment. As illustrated above (table 1), Roma are heavily concentrated in the lower ends of the educational attainment scale, whilst, non-Roma on average have much higher levels of attainment and are almost absent from the lowest 'no education' level. The approach used here explicitly takes this issue into account.

Specifically, I employ the method proposed by Ñopo (2008) to gender and ethnic wage gaps in order to identify explained and unexplained characteristics. This involves person-to-person matching (with re-sampling) which bases the estimates of explained and unexplained components

on observed differences in outcomes for which there is common support[3]. More precisely, the approach involves one-to-many 'perfect' matching. A female is selected from the sample and her wages are compared to the average of all men in the sample who apart from their gender possess the same characteristics[4]. The process is repeated (with replacement in the male sample) until the entire sample of interest has been considered. At the end some males and some females may not have found matches due to their lack of shared characteristics; these two groups are outside the common support. As a consequence, the mean difference in male/female outcomes can be decomposed into three explained and one 'unexplained' components comprising[5]:

a. Differences arising between women for whom there is common support and those for whom there is not;
b. Differences arising between men for whom there is common support and those for whom there is not;
c. Differences arising between women and men within the range of common support due to differences in their characteristics; and,
d. Differences which cannot be explained by any of the three elements a. - c. above – the unexplained component.

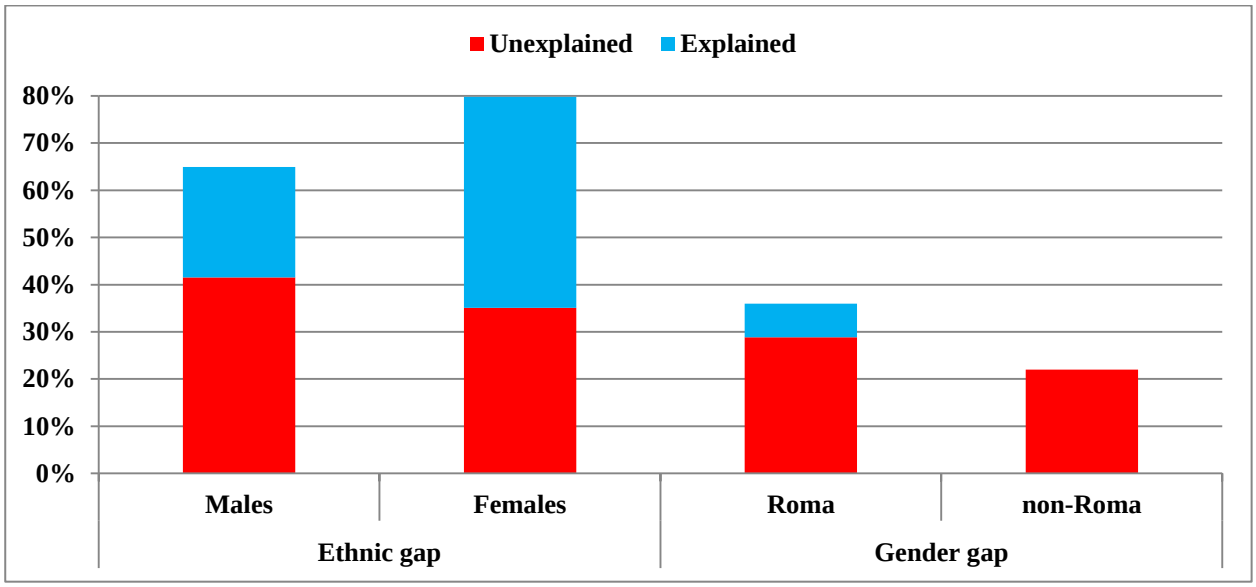
This last element provides an estimate of 'discrimination' analogous to that estimated parametrically with the Blinder-Oaxaca type method, however, it is explicitly based only on those individuals who possess the same characteristics; for whom there is 'common support'. The method explicitly limits comparisons to those with the same characteristics (other than gender), making no assumptions about effects outside the field of observation; the estimate of discrimination is based on the notion that those with the same characteristics (apart from gender) should have the same wages.

[3] It is precisely this characteristic which makes the methodology useful here. One may observe that also other potential matching approaches – such as propensity score matching which is otherwise an obvious alternative candidate for use in this type of exercise, is also extremely susceptible to the failure of common support since persons with similar propensity scores may have - and in this case almost certainly will have – quite different (educational) characteristics.
[4] For simplicity, the methodology is presented in terms of the identification of gender based discrimination, however, in what follows we look at both gender and ethnic based differences using the same methodology.
[5] For a more complete - and technical - explanation, the reader is referred to the original Ñopo (2008) article.

The method also allows a simple examination of unexplained differences across individual characteristics which allow some analysis of the factors driving gender wage gaps. The major drawback with the method is the so-called 'curse of dimensionality'. In common with other non-parametric and semi-parametric approaches, as the number of characteristics forming the basis of the 'common support' is increased the number of matches is correspondingly reduced – reducing the field of common support. This essentially means that, the number of base characteristics – or controls – must be relatively limited. Specifically, controls for country, education, (potential) experience are added in addition to gender and, for the first part of the analysis, ethnicity.

The issue is approached in two ways. First, ethnic based differences are examined separately for men and women. This allows a comparison of the degree of ethnic differences in the unexplained wage gaps between men and women. Second, the issue of gender based differences wage gaps are analysed directly. Figure 5 reports the results of these estimations with the ethnicity based comparison on the right and gender on the left. The height of the bars in the figure represent the relevant (ethnic or gender based) wage gap as a percentage of the wage of the group who are 'disadvantaged' (Roma on the left and women on the right); thus it illustrates the fact that the size of the ethnicity based gap is larger than the gender based one and that gap is larger for women (and correspondingly the gender gap is larger for Roma). The content of the bars in the left of the figure show that the lower educational attainment of Roma explains a significant portion of the Roma/non-Roma wage gap and that this explained portion is larger for women (56% of the gap) than men (36%). In consequence, the unexplained portion attributable to other factors including labour market discrimination is – expressed as a percentage of the gender specific Roma wage – roughly similar for men (42%) and women (35%). The right hand side of the figure shows clearly that only a very small portion (20%) of the gender gap for Roma – and none for non-Roma – is explained by educational differences.

Figure 5: Estimation of explained and unexplained differences in wage gaps using nonparametric matching



Source: estimated on the basis of the UNDP/WB/EC regional Roma survey 2011.

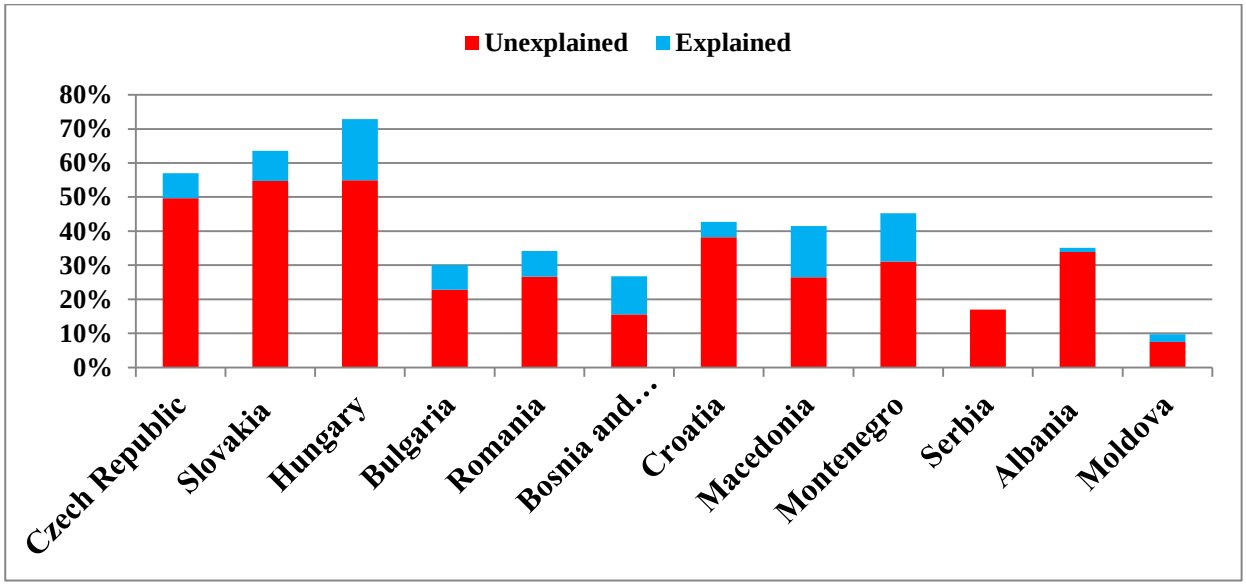
Notes: 1) the figure reports the results of estimating the explained and unexplained components of wages using the non-parametric matching technique proposed by Ñopo (2008).
2) the height of each bar is the gap between the mean wage of Roma and non-Roma (left) or Males and Females (right) expressed as a percentage of the mean wage of the 'disadvantaged' group (Roma or women).
3) On one occasion the estimate of the 'unexplained' portion was above 100% of the total gap; in this case the 'unexplained portion was set to 100%.
4) more detailed numerical results are reported in the appendix.

Confining attention to the Roma, and undertaking the Ñopo decomposition on a country by country basis (figure 6), one can observe that both the gender gap and the proportion of it which is explained by education varies significantly across countries. The percentage of the gender wage gap explained by differences in educational attainment and experience varies from zero (in Serbia) to 42% (in Bosnia and Herzegovina). Nowhere however, does it explain anything like one half of the wage gap[6]. Interesting too, is that fact that the earliest entrants to the EU in this group – Czech Republic, Hungary and Slovakia, are precisely the countries where the gender wage gap is greatest. Whilst in one sense it is encouraging that differences in educational attainment count for very little in these countries – as in the newer EU Member States, Bulgaria and Romania as well as the most

[6] To compare, the percentage of the ethnicity based gap explained by education and age reached 76% (in Albania) and was above 50% for women in six countries and for men in three. See the appendix.

recent entrant, Croatia – the fact that a very substantial unexplained gender based wage gap exists is, or should be, a matter of some concern.

Figure 6: **Estimation of explained and unexplained differences in gender wage gaps using nonparametric matching**



Source: estimated on the basis of the UNDP/WB/EC regional Roma survey 2011.

Notes: 1) the figure reports the results of estimating the explained and unexplained components of wages using the non-parametric matching technique proposed by Ñopo (2008).
2) the height of each bar is the gap between the mean wage of men and women expressed as a percentage of the mean wage of the 'disadvantaged' group (Roma or women).
3) On occasion the estimate of the 'unexplained' portion was above 100% of the total gap; in these cases the 'unexplained portion was set to 100%.
4) more detailed numerical results are reported in the appendix.

In any event, controlling for educational attainment does little to reduce the gender wage gap so some other source for it must be sought. A number of potential candidates offer themselves for consideration. Marriage and child-bearing, and more generally 'traditional' social norms might limit the labour market opportunities of Roma. On the other hand, occupational and industrial segregation might also be playing a role. The distinction – indeed the relative importance of these different types of factors - is important since in the first case one is primarily making reference to models of behaviour which are imposed internally in the Roma community whilst the latter can be – arguably – more directly related to discriminatory factors in the labour market. Although an

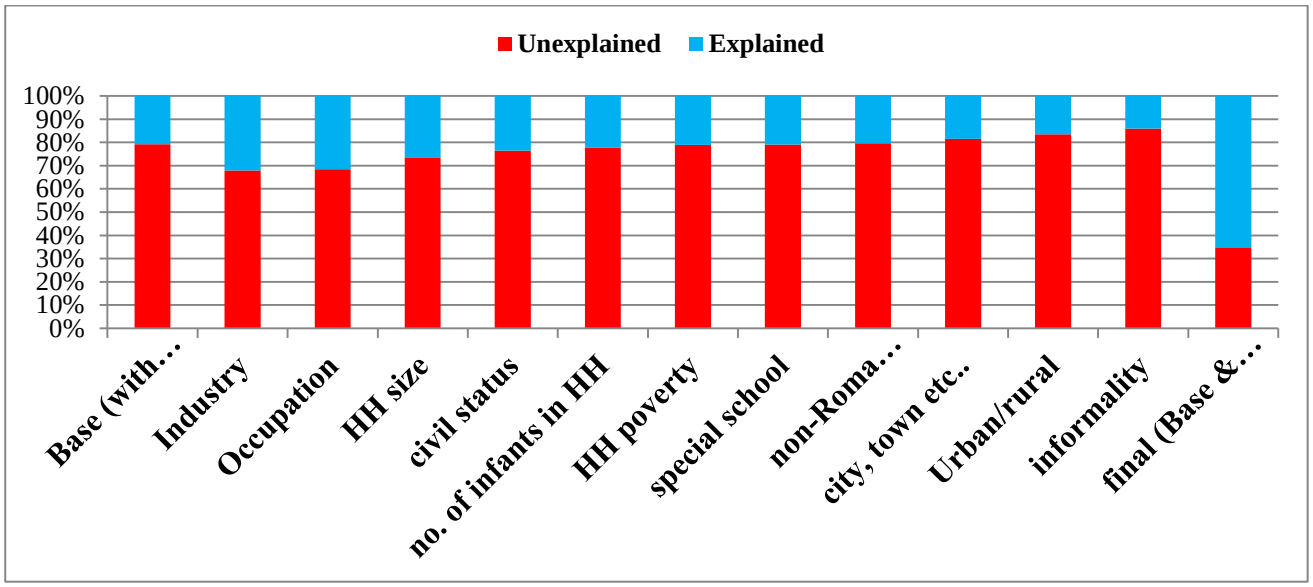
oversimplification – also the type of job open to Roma women may depend in part on community social norms – identifying exactly where the problem lies may provide a basisi on which to look for solutions. I consider a variety of factors where which might contribute to the wage gap: household size, primarily interpreted as an indicator of socio-economic status; the number of young children in the household; urban or rural environments; whether the person lives in a city, town, village or unregulated area; the industrial and occupational classification of the job in order to identify elements of segregation; informality of employment – a significant factor in explaining the 'unexplained' portion of Roma/non-Roma employment and wage gaps (Kahanec and Yuksel, 2010; O'Higgins, 2011); whether the individual attended a special school – as noted above, an area which has been investigated as a possible explanatory factor in the ethnic wage gap in the Czech Republic and Slovakia (O'Higgins & Brueggemann, 2013); and, whether the household in which the person lives is poor and so on.

Figure 7 summarises the results. The estimation procedure added each variable to the base specification in turn and then combined those with the greatest explanatory power into the final form shown by the bar furthest to the right of the figure[7]. Between the base and final specification, the bars are ordered by their individual 'explanatory power' that is by the extent to which each one taken singly reduces the unexplained portion of the wage gap when added to the base specification. On their own, industrial sector and the type of occupation are the most important factors in explaining the previously unexplained portion of the gender wage gap – each of these reduces the unexplained portion of the gap by around 11 percentage points[8].

[7] Starting with the variable with the largest explanatory power, measured by the reduction in the unexplained portion of the wage gap produced the inclusion of the variable on its own, variables were added to the specification and kept if they lead to a further reduction in the unexplained portion of the wage gap also ion the presence of other explanatory factors. Thus, for example, neither the direct measure of household poverty nor the number of infants in the household are included in the final specification since, once household size was included, these factors did not add to the explanatory power of the specification. It might be added that, in contrast to the more traditional parametric approach, it is by no means automatic that adding variables improves the explanatory power of a specification.

[8] It might be observed that, given the overlap between industrial sector and occupation, taken together these two factors 'only' add 15 percentage points to the explained portion of the wage gap.

Figure 7: **Estimation of explained and unexplained differences in gender wage gaps with differing control variables**



Source: estimated on the basis of the UNDP/WB/EC regional Roma survey 2011.

Notes: 1) the figure reports the results of estimating the explained and unexplained components of wages using the non-parametric matching technique proposed by Ñopo (2008).
2) since the base gap is always the same, it is set here to 100 so that explained and unexplained portions of the wage gap are here expressed as a percentage of the total gender gap.
3) the bar furthest to the left reports the base specification; that is, including controls for country, age and education, that furthest to the right reports the final specification including, in addition to the variables in the base specification, industry, occupation, household size, civil status and whether or not the individual lives in a predominantly Roma community. The other bars report the base specification with the addition of the single indicator specified below it.
4) more detailed numerical results are reported in the appendix.

Thus, there are strong grounds to support the notion that both industrial and occupation segregation are playing a key role in determining the Roma gender wage gap. Other factors which proved important were household size, an individual's civil status, and whether the individual lived in a predominantly non-Roma community or not. I would suggest that household size is acting as a proxy for the socioeconomic status of the individual – an interpretation that is supported by the fact that once this variable is included, neither the number of infants in the household nor the direct (expenditure-based) measure of household poverty adds explanatory power to the model. The role of an individual's civil status in the gap may, I would suggest, be explained by differential effects that marriage and cohabitation will have on Roma men and women's employment and earnings, with the male being the primary breadwinner and the female providing, in normal circumstances a

secondary source of income. Finally, living in a predominantly Roma community (or not) also appears to be an important part of the explanation. In the final specification, around 65% of the wage gap is explained by these factors – in addition to education and experience.

Major factors underlying the wage gap have been identified, however, although the analysis gives an idea of the relative importance of different factors, it does not say anything about the direction of the effects – for example, is the gender gap larger in Roma dominated or non-Roma dominated communities? On this, the non-parametric approach remains silent, so one is forced albeit somewhat reluctantly to return to a more traditional albeit simple econometric specification to further investigate this question. However, I hasten to point out that no causal interpretation should be attached to this simple model[9]. The point is primarily to identify the sign of the association between the wage gap and the 'explanatory' variables.

Regression of the natural logarithm of monthly wage on the factors identified above undertaken separately for men and women allows us to throw some light on this issue (table 2). The regression includes a full set of country, industry and occupation fixed effects not reported in the table. Once occupation and industry are controlled for, educational attainment is an important explanatory factor and indeed, the returns to education are greater for women than men[10]. One would not wish to attach too much significance to this observation; as noted above, the educational level itself is likely to be endogenous, as our occupational and industrial choices. The main point of the regression lies in the examination of the other coefficients in order to throw light on the role of these factors in explaining the Roma gender wage gap.

[9] There are several reasons for this; not least because most – if not all – the 'explanatory' variables may be endogenous. This explains too why no attempt is made to control for sample selection bias. As I have argued elsewhere (O'Higgins, 2010) such an approach is inadequate in the current context which would require a more complete simultaneous model.

[10] Although it is fairly evident from the reported standard errors and has been confirmed by joint estimation not reported here, that the gender coefficient difference is not statistically significant. One might also observe, however, that excluding occupational and industrial dummy leads to larger returns to education for both women and men whilst the greater returns to education for women – compared to men - is maintained

Table 2: OLS wage regression for Roma by gender

	Males		Females	
	Coeff	**s.e.**	**coeff**	**s.e.**
Educational attainment:				
- primary	0.10**	0.041	0.18***	0.058
- lower secondary	0.26***	0.041	0.36***	0.059
- upper secondary	0.35***	0.050	0.48***	0.076
- post-secondary	0.52**	0.201	0.76***	0.201
Experience/10	0.17***	0.061	0.05	0.08
Experience squared/1000	-0.36***	0.102	-0.12	0.138
Ln (HouseHold size)	-0.03	0.029	-0.06	0.041
Married	0.14***	0.041	0.02	0.043
Roma community	-0.01	0.030	0.11***	0.040
adjusted R-squared	0.290		0.347	
N	3092		1372	

Source: estimated on the basis of the UNDP/WB/EC regional Roma survey 2011.

First, the (natural logarithm of) household size[11] has a negative sign which is larger for women although in both cases this is not statistically significant. Being married does appear to have a strongly differentiated effect. For men, being married is associated with significantly higher wages, whereas no such association is present for women. There are good reasons to suppose that the causal effect is not unidirectional – quite possibly a higher wage makes marriage more likely for men for example, however, for the purposes here, the suggestion is that the higher wages – ceteris paribus – of married men is driving the role of marriage in the non-parametric estimation procedure above rather than the lower wages of married (compared to unmarried) women. Perhaps most interestingly, living in a predominantly Roma community is associated with higher wages for Roma women but not so for Roma men. I have no explanation for this, however, it would certainly bear further investigation.

[11] Specifically, the number of members of the household.

7. Conclusions

A number of studies which have looked at the issue have found that Roma/non-Roma wage and employment gaps cannot be explained entirely by the lower levels of educational attainment of Roma. For the most part, these have been based of parametric estimation procedures using some form of linear or non-linear Blinder-Oaxaca type decomposition of the gaps into explained and unexplained components. The unexplained component is then typically attributed to labour market discrimination. Although the multiple disadvantages faced by Roma women has been noted by a few authors, there have not yet been any studies – to my knowledge – which have looked at the question systemically, still less tried to explain the source of this disadvantage.

This paper has then several purposes, in the first place to suggest a more appropriate framework for the analysis of Roma wage and employment gaps. The parametric approach adopted by the bulk of the literature is based on a number of assumptions but of most relevance here, the existence of 'common support' that is that Roma and non-Roma individuals have similar characteristics. Since it is precisely differences in the level educational attainment which distinguish Roma from non-Roma, the adoption of an approach which assumes similarity between them is arguably inappropriate. A more appropriate non-parametric method suggested by Ñopo (2008), and that adopted here, makes no such assumption and confines explicitly confines comparisons to individuals with the same characteristics.

The second purpose has been to quantify the size of the gender gap amongst women and see the extent to which the lower educational attainment of Roma women (compared to men) may explain this difference. The analysis has shown that the Roma/non-Roma wage gap is larger for women than for men – or, to put this another way, the gender wage gap is larger for Roma than non-Roma - although a greater portion of this gap is explained by differences in educational attainment between Roma and non-Roma women (compared to Roma and non-Roma men). Looking at the gender gap for Roma explicitly, it was found that on average –across the 12 countries studied here –

around one-fifth of this gap can be explained by differences in educational attainment between Roma men and Roma women.

The paper then proceeded to seek possible sources for the remainder of this gap. Differences in occupation and industry account for a further 15% of the gap. This is important in that it suggests that explanations for gender differences in pay amongst Roma need also to look at occupational (and industrial) segregation issues. The suggestion is that the observed 'labour market discrimination' is more a question of 'differential pay for work of equal value' than 'different pay for the similar work'. Other important factors in explain the gender wage gap were found in household size, which I would suggest is primarily acting as a proxy for household socio-economic status; an interpretation which is further supported by the fact that when household size is included, other variables representing household income (household poverty and the number of infants in the household) which on their own influence the gender gap, lose all relevance when a household size variable is included. A similar point may be made concerning the role of living in a predominantly Roma community. This clearly dominates other relevant community characteristics such as the size of the settlement (city, town, village etc.,) or whether the area is rural or urban. The final regression analysis threw light on this. It appears that Roma women living in predominantly Roma communities earn more than similar women living in communities where non-Roma are in a majority. I offer no explanation for this, however, it is likely that this may have to do with country specific factors, and I suggest would be worthy of further investigation.

Finally, with the inclusion of a variety of characteristics a substantial portion of the Roma gender wage differential - some 35% - remains unexplained. In the absence of any other competing explanations, this may reasonably attributed to the additional labour market discrimination faced by Roma women when compared with their male counterparts.

References

Kahanec, M. and Yuksel, M. 2010. "Intergenerational Transfer of Human Capital under Post-War Distress: The Displaced and the Roma in the Former Yugoslavia." in G. S. Epstein, I. N. Gang (ed.s) *Migration and Culture*, (Bingley, Emerald Group Publishing Limited), pp.415-443.

Kling, J. and Kagin, J. 2012. 'Integrated household surveys among Roma populations: one possible approach to sampling used in the UNDP-World Bank-EC Regional Roma Survey 2011', Roma Inclusion Working Papers. Bratislava, UNDP.

Kosko, S. 2012. Educational attainment and School-to-work conversion of Roma in Romania: adapting to feasible means or ends? *Journal of Human Development and Capabilities,* Volume 13, no. 3, 415-50.

Milcher, S. 2011. Decomposing income differentials between Roma and non-Roma in South East Europe. *Romanian Journal of Regional Science*, 5, no. 1, 27–53.

Milcher, S. and Fischer, M.M. 2011. On labour market discrimination against Roma in South East Europe, *Papers in Regional Science*, 90, no. 4, 773-88.

Ñopo, H. 2008. "Matching as a tool to decompose wage gaps," *Review of Economics and Statistics*, Vol. 90, no. 2, 290-9.

O'Higgins, N. 2010. "Its not that I'm a racist, its that they are Roma: Roma Discrimination and Returns to Education in South Eastern Europe. *International Journal of Manpower*, 31, no. 2, 163-87.

O'Higgins, N. 2011. 'Roma and Employment in Macedonia', Research Report, Skopje, UNDP.

O'Higgins, N. 2012. 'Roma and non-Roma in the Labour Market in Central and South Eastern Europe', Roma Inclusion Working Papers, UNDP, Bratislava.

O'Higgins, N. and Brüggemann, C. 2013. 'The influence of educational attainment, special schooling and labour market discrimination on employment of Roma in Central Europe', Paper presented at the annual meeting of the American Educational Research Association (AERA), University of San Francisco.

Revenga, A., Ringold, D. and Tracy, W. M. 2002. *Poverty and ethnicity: A cross-country study of Roma poverty in Central Europe*, Washington DC, World Bank.

Trentini, C. 2011. Ethnic patterns of returns to education in Bulgaria: Do minorities have an incentive to invest in education? UNECE Discussion paper series no. 2011/1, Geneva, UNECE.

Appendix – Additional tables

Table A1: Nopo decomposition of ethnic gaps in wages by country and gender (underlying figure 5 in the text).

	Males			Females		
	Total	**unexplained**	**Std. err. unexplained**	**total**	**unexplained**	**Std. err. unexplained**
All countries	0.65	0.42	0.02	0.80	0.35	0.02
albania	0.52	0.18	0.04	0.63	0.15	0.06
bosnia and herzegovina	0.92	0.51	0.06	1.39	0.91	0.12
bulgaria	0.67	0.44	0.09	0.61	0.50	0.02
czech republic	0.60	0.34	0.04	0.37	0.10	0.03
slovakia	0.33	0.12	0.07	0.70	0.26	0.16
Montenegro	0.33	0.13	0.16	1.20	0.92	.
croatia	0.58	1.24	.	0.80	0.37	.
hungary	0.48	0.35	0.04	0.34	0.20	0.07
macedonia	0.71	0.55	0.05	0.96	0.58	.
moldova	0.09	0.05	0.28	0.28	0.38	0.16
romania	0.69	0.42	0.10	0.84	0.37	0.07
serbia	0.71	0.36	0.07	1.16	0.55	0.02

Note: . indicates the standard deviation was not calculable due to the lack of sufficient common support.

Table A2: Nopo decomposition of gender wage gaps for Roma by country (underlying figures 5 & 6 in the text).

	total	unexplained	Std. err. unexplained
All countries	**0.36**	**0.29**	**0.03**
Albania	0.35	0.34	0.05
Bosnia and Herzegovina	0.27	0.16	0.12
Bulgaria	0.30	0.23	0.06
Croatia	0.43	0.38	0.05
Czech Republic	0.57	0.50	0.10
Hungary	0.73	0.55	0.28
Macedonia	0.42	0.26	0.12
Moldova	0.10	0.07	0.07
Montenegro	0.45	0.31	0.09
Romania	0.34	0.27	0.13
Serbia	0.17	0.27	0.10
Slovakia	0.64	0.55	0.10

Table A3: Nopo decomposition of gender wage gaps for Roma, specification search (underlying figure 7 in the text).

	total	**unexplained**	**Std. err. unexplained**
Base (with education, age & country)	**.36**	**0.29**	**.03**
Industry	.36	0.25	.02
Occupation	.36	0.25	.03
HH size	.36	0.27	.03
civil status	.36	0.28	.04
no. of infants in HH	.36	0.28	.03
HH poverty	.36	0.29	.03
special school	.36	0.29	.03
Roma community	.36	0.29	.03
city, town etc..	.36	0.30	.03
Urban/rural	.36	0.30	.03
Informality	.36	0.31	.03
Final (Base + industry, occupation, HH size, civil status and Roma community)	**.36**	**0.13**	**.03**