

Lehmer, Florian; Ludsteck, Johannes

Conference Paper

The Immigrant Wage Gap in Germany: Are East Europeans Worse Off?

50th Congress of the European Regional Science Association: "Sustainable Regional Growth and Development in the Creative Knowledge Economy", 19-23 August 2010, Jönköping, Sweden

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Lehmer, Florian; Ludsteck, Johannes (2010) : The Immigrant Wage Gap in Germany: Are East Europeans Worse Off?, 50th Congress of the European Regional Science Association: "Sustainable Regional Growth and Development in the Creative Knowledge Economy", 19-23 August 2010, Jönköping, Sweden, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/118977>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

The Immigrant Wage Gap in Germany: Are East Europeans Worse Off?

Florian LEHMER #
Johannes LUDSTECK

Institute for Employment Research Nuremberg
April 2010

Abstract

This study compares the outcomes of male foreign workers from different East and West European countries who entered the German labour market between 1995 and 2000, with those of male German workers. We find that the immigrant-native wage gap differs significantly between nationalities: the differential is largest for workers from Poland (-44 percent) and the Czech Republic (-38 percent) and by far the lowest for Spaniards (-8 percent). Results from an Oaxaca/Blinder type decomposition show that unfavourable characteristics (compared with German workers) contribute significantly to the explanation of the immigrant wage gap. This is especially true for workers from Poland, Portugal, Italy and Slovakia. For all other countries, it is observed that the coefficients effect dominates. It can therefore be concluded, that immigrants are generally affected by “discrimination”. Comparing the effects for workers from East European EU member countries with those for other nationality groups, it emerges that East Europeans are not worse off than other nationalities. The most pronounced “discrimination” is found for immigrants from non-EU states in Eastern Europe.

To analyse the importance of segregation into sectors, we take a closer look at construction and hotels & restaurants and find that the coefficients effect still adds most to the explanation of the raw wage differential between foreigners and Germans. This indicates that segregation into sectors does not significantly contribute to the “discrimination” of foreigners.

Additional information is obtained from quantile decompositions. Coefficient effects (in absolute values) decrease for the majority of countries. Thus, discrimination appears to be more pronounced at low wage levels. Moreover, this evidence suggests sticky floors rather than glass ceilings.

Keywords: East Europeans, immigration, wage gap, (quantile) decomposition, EU enlargement, sticky floors.

JEL classification: J61, J31, J15.

Correspondence to: Dr. Florian Lehmer, Institute for Employment Research Nuremberg, Regensburger Straße 104, D-90478 Nuremberg, Germany, e-mail: florian.lehmer@iab.de, phone: +49 (911) 179-5664, fax: +49 (911) 179-3296.

1 Introduction

In Germany the eastern expansion of the EU has led to fears that the labour market will be congested by workers from the new member countries. Due to large wage differentials between Germany and Eastern Europe, especially low-skilled workers worry that they may be substituted by East European workers. In order to allay such worries, immigration from East European EU member states is still to be regulated until 2011. Though some studies (see e.g. Baas et al., 2007) try to assess the effect of immigration from East European countries on wages in German regions, there is no empirical study comparing the wages of immigrants from the new EU member countries with those of the native population and migrants from other countries. Due to sample size restrictions, existing studies on immigrants wage gaps have to pool immigrants from different countries, thus neglecting possibly noteworthy differences between them which are of importance in economic terms.

This paper seeks to close these research gaps. Considering workers from Eastern Europe who entered the German labour market between 1995 and 2000, we first investigate the characteristics of this group. More specifically, we are interested in the skill composition and the industries in which they start to work. Using a rich data set we examine this issue in depth by observing individuals from different East European countries and comparing them with other ethnic groups (from different West and South European countries). Second, we are interested in the immigrants' outcomes. Especially with regard to the full opening of the German labour market to East Europeans in 2011, we ask whether East Europeans face particular disadvantages (in terms of wages) in the labour market or not. Some of the East European immigrants, for instance Bulgarians, Hungarians or Slovaks, show higher formal qualification levels on average than immigrants from other EU countries. The qualification level of other East European immigrants (for instance Poles and Czechs), however, is below average. Considering that considerable immigration from the countries of Eastern Europe took place in the last decade despite tough entry quotas, East European immigrants appear to be a highly motivated group. To compare this group with immigrants from other countries, we measure the earnings gap between Germans and immigrants separately for each nationality and decompose it into a characteristics (endowment) effect and an unexplained (coefficients) effect. Under assumptions explained in the

estimates section below, the latter effect can be interpreted as a measure of discrimination. Our cross-country comparison of the decomposition effects shows whether some nationalities are affected by discrimination more strongly than others. To the best of our knowledge, there is no other study for Germany that considers the immigrant wage gap in such depth.¹ The large size of our sample allows us to examine the issue even more closely: nuanced views of discrimination distinguish between markets where the remuneration of immigrants' productive endowments is lower (pure discrimination) and markets where immigrants obtain lower wages because they sort themselves (or are sorted) into low-wage industries (segregation). To assess the contribution of sorting to the gross wage gap, we perform decompositions for specific industries with substantial immigrant worker shares (*construction; hotels & restaurants*) and compare them with the pooled (all industries) base sample.

Moreover, we apply quantile decomposition techniques to examine the heterogeneity of discrimination and endowment effects across the wage distribution. This may be of importance particularly since the characteristics (mainly the qualification level) of immigrants from some countries are more homogenous than others and this may have important effects on the distribution of the wage gap.

As in many other countries, there is overwhelming evidence of the existence of an immigrant-native wage gap in Germany. For instance, Diekmann et al. (1993) find that foreign male workers earned 9 percent less than German male in 1985². Aldashev et al. (2008) obtain a raw male foreigner wage gap of 11 percent from the GSOEP survey. According to the authors, less than half of this gap can be explained by differences in endowments, leaving considerable scope for discrimination. However, since foreigners are broadly defined as non-Germans (due to sample size restrictions), this study provides no information on nationality-specific differences within the groups of foreigners.

Besides differences in endowments (mostly qualification levels) and discrimination, a further explanation is provided by the "assimilation" literature dating back to Chiswick (1978). This tries to explain the wage gap by the fact that human capital is specific to the host country. With time spent in the host country, immigrants acquire language skills, accumulate other general human capital and become acquainted with the host country's labour market. Through this assimilation process,

¹ A similar analysis was conducted by Nielsen et al. (2004) for Denmark. These authors observe the immigrant wage gap separately for workers from Nordic countries, Turkey, Africa, Pakistan and India, and Sri-Lanka.

² Their results are based on the Mikrozensus 1985.

immigrants should typically be seen to catch up with the native workers' wages.³ This issue may be crucial especially when East European immigrants are compared with other immigrants since they a) have considerably shorter experience on average in the German labour market and b) have almost no access to networks of compatriots already living in Germany. The large size of our sample allows us to improve the comparability between the East European immigrants and immigrants from other countries by restricting the sample to cohorts entering Germany between 1995 and 2000.

The remainder of the paper is organised as follows: the next section deals with a description of our data source and presents some basic information on nationality-specific differences in characteristics. Section 3 describes the estimation approach and presents the results. Section 4 concludes.

2 Data and descriptives

We use the employment register data (BEH) of the German Federal Employment Agency for the period 1995-2006. Its crucial advantage for our application is its size: it covers nearly 80 percent of the German workforce, excluding only the self-employed, civil servants, individuals in (compulsory) military service, and individuals in so-called 'marginal part-time jobs' (jobs with no more than 15 hours per week or temporary jobs that last no longer than 6 weeks).⁴ It also contains important personal characteristics (sex, age, qualification level, job status) as well as information on occupation, industry, establishment identifiers, wages, and regional information which refers to both the location of the firm/workplace and the place of residence at NUTS3 (district) level. The nationality variable in the data is of particular interest for our analysis.⁵ We select individuals from "classical" EU countries (Greece, Portugal, Italy and Spain), from East European EU member states (Romania, Poland, Hungary, Czech Republic, Slovakia and Bulgaria), other East European countries (Ukraine, Belarus and Russia) and Turkey. *Table 1* shows the observation figures for different nationalities. For instance, the employment register data of the period 1995-2006 includes 1.3 million observations of Greeks (see

³, This is corroborated empirically by Borjas (1987) for the US.

⁴ For a detailed description of the data set see Bender et al. (2000) or Bender et al. (1996). A more commonly used data set in Germany is the IABS, which is a 2 percent random sample of the data set we use.

⁵ Note, however, that German resettlers who immigrated from the (former) Soviet Union are Germans by law when immigrating and are therefore not included in the foreigner sample.

column A95/06 in *table 1*). Concentrating on the more recent period of 2001 to 2006 there are still about 650,000 observations (column A01/06). The sheer size of the data source affords the opportunity to increase the comparability between immigrants from different countries to a large extent: we select a cohort of individuals entering the German labour market for the first time between 1995 and 2000 and for whom we are able to obtain at least one valid wage observation between 2001 and 2006.⁶ It can be seen from *table 1* that the number of observations decreases substantially as a result of this restriction (see column E01/06). Moreover, due to the different labour market situations in western and eastern Germany we restrict our analysis to western Germany (see column E01/06W).⁷

[*Table 1* about here]

Although conditioning on the entry years 1995-2000 eliminates some of the differences between the 14 groups with regard to labour market assimilation⁸ and age effects, major differences remain with respect to gender, employment type and working time. These differences are most pronounced between the “classical” countries of origin Italy, Greece, Spain, Portugal and Turkey on the one hand and the East European countries on the other hand. It is evident from *table 2* that women are clearly over-represented in the groups of immigrants from Bulgaria (67 percent), Hungary (63 percent), Poland (65 percent), Romania (66 percent) and Slovakia (60 percent), when this is compared with the 48 percent share of German female workers. In contrast, in the immigrant groups from Italy, Greece, Portugal and Turkey women are clearly under-represented (between 44 percent and 47 percent). Interestingly this does not apply to Spanish immigrants, where the share of women is comparatively large (56 percent), as it is for the East European sample.⁹ Due to the substantial differences in the gender distribution and in order to further increase the homogeneity of the selected sample, we exclude female workers from the analysis in this paper. Moreover, we exclude apprentices and restrict

⁶ More specifically, we computed the entry years (defined as the first appearance of the employee in our reference date data sets) for all foreigners and dropped all individuals with an entry year before 1995 or after 2000. Without this restriction, our results would potentially be biased since it is reasonable to assume that immigrant cohorts differ with respect to their characteristics and that there is a time trend in the productivity of immigrant cohorts (for further details, see, for instance, Lalonde and Topel, 1997).

⁷ More specifically, we check the employment history of all of the workers in the sample and exclude everyone who had ever worked in eastern Germany.

⁸ As pointed out above, labour market participation should have a positive impact on the immigrant’s productivity and wages (see, for instance, Chiswick, 1978 or Chiswick and Miller, 2007).

⁹ The remaining countries (Russia, Czech Republic, Ukraine and Belarus) also reveal minor differences compared with the gender distribution of Germans.

the sample to full-time workers aged between 25 and 55, since it emerges that - even after dropping the female workers - there are still major differences with respect to employment type (full-time, part-time, apprenticeship) and working time.¹⁰ This final sample is labelled S01/06W in *table 1*.¹¹

[*Table 2* about here]

After this restriction to full-time male workers we present some evidence on nationality-specific differences with respect to the qualification level and industry affiliation. These two dimensions probably explain the lion's share of the wage differential between nationalities due to individual-specific differences in characteristics.¹² *Table 3* compares the proportions of low-skilled, skilled and highly-skilled individuals for each nationality.¹³ As a further category we include *skill missing* in order to account for systematic differences between Germans and foreigners regarding the reliability of the skill variable. It is evident that information on the skill level is missing for 10 percent of German workers whereas the corresponding values lie between 16 percent for Czechs and more than 30 percent for Poles and Portuguese nationals. Regarding the observable skill levels, it can be seen that foreign workers are generally over-represented in the group of low-skilled workers. The highest values are obtained for Greeks (39 percent), Italians (33 percent), Portuguese (37 percent) and Turks (39 percent), and the lowest values for Bulgarians, Slovaks (both 17 percent) and Hungarians (14 percent). Turning to the highest skill level, it can be observed that 13 percent of all male German workers come in this category. There are marked differences between the groups of foreign workers. Italians and Poles (each about 5 percent), Portuguese and Turks (each about 2 percent) are clearly under-represented in this category, the opposite is true of Bulgarians (29 percent), Spaniards (23 percent) and

¹⁰ Since working time is only reported in three classes, this restriction additionally avoids a potential bias due to imprecise working time information. To keep the paper size small, we do not present descriptive evidence on nationality-specific differences in employment type; but it is available from the authors on request.

¹¹ As a final restriction, the study is limited to individuals with reliable wages. Wages which are calculated as a daily average over the observed employment period for each person are presumed to be unreliable if they are below a specific level. We take double the minimum income threshold for compulsory social insurance in a given year as this level.

¹² A further possible important source of wage differentials is the region (and region type) of destination. It emerges that immigrants predominantly decide to settle in the metropolitan areas of North Rhine-Westphalia, Bavaria and Baden-Württemberg (again the results are not included in the paper but are available on request). Czechs who work in the rural areas in the eastern part of Bavaria are exceptions. Generally, it can be assumed that the location decision strongly depends on (i) given earnings and employment opportunities in a region, (ii) the distance from the home country and (iii) the existence of locational networks of the specific nationality group. The latter argument is analysed in depth, for instance, by Bartel (1989), Zavodny (1999), Bauer et al. (2002) and Bauer et al. (2005).

¹³ A description of the variables is provided in Table A1 in the Appendix.

Ukrainians (21 percent). In the intermediate category, the values for foreign workers are below the share of Germans in each case. Altogether this indicates that there are notable skill-specific differences between Germans and immigrants. Moreover, this is a first indication that immigrants can not be regarded as a homogenous group, (as is the case in other studies on immigrants). Hence, the specific characteristics of each nationality should be dealt with separately. In addition to this, the reporting of the qualification variable in the register data is probably less reliable for immigrants, as employers (who have to fill in the notification forms for the social security register) may often be unsure how qualifications gained in foreign countries are comparable with their German counterparts. Furthermore, reporting seems to be biased toward the task performed, i.e. employers sometimes report the qualification level required by the job performed instead of the true qualification level. This implies that qualification levels are under-reported especially for immigrants as they are more often overqualified for their work than native workers. In the appendix we provide a loose robustness check for this by restricting the estimation sample to the medium-skilled where the reporting of qualification levels can be expected to be more reliable than for the whole sample.

[Table 3 about here]

The heterogeneity of foreigners is also evident when comparing industry affiliations. *Table 4* presents the three most important industries for each nationality together with the industry-specific median wage of each nationality.¹⁴ The top industries for German male full-time workers are *construction* (9 percent), *other business activities* and *machinery & equipment* (both 7 percent). Though *construction* and *other business activities* generally play an important role for most of the foreign nationality groups too, we observe huge differences with respect to the degree of segregation into branches. For instance, more than a quarter of the Portuguese immigrants but only few Bulgarians or Ukrainians work in the *construction* sector. The latter are primarily employed in *other business activities* (14 percent), almost double the share of German workers in this sector.¹⁵ This segregation is also obvious in the *hotels &*

¹⁴ More specifically, we present the results for the two-digit industry classification (WZ 2003) of the Federal Statistical Office.

¹⁵ In our classification, this sector is very heterogeneous and includes, for instance, high-wage industries like consulting, legal advice, architecture or advertising and low-wage industries like temporary agency work, security services and cleaning services. A closer look at the data shows that the employment figures are clearly higher in the low-wage industries, indicating that foreigners are overwhelmingly employed as temporary workers or cleaning staff.

restaurants sector. 24 percent of Slovaks, 17 percent of Italians and 18 percent of Bulgarians are employed in this sector whereas it is somewhat less relevant for Greeks (11 percent), Hungarians (12 percent), Portuguese (10 percent), Romanians (7 percent) and Spaniards (6 percent). In contrast, this sector is not relevant at all for Poles, Russians, Ukrainians and Belarusians (or for Germans),¹⁶ Further peculiarities are that Ukrainians choose to work in *health & social work* (7 percent), while Poles strongly select themselves into the *agricultural* sector (23 percent)¹⁷, or that Bulgarians (8 percent) and Poles (6 percent) have high preferences for working in the *recreation* sector. To sum up, it is evident that foreigners are not evenly distributed across the same industries as Germans. It is more the case that they sort themselves into specific branches which differ across the foreign nationality groups.

Besides the selection into industries we also observe differing remuneration within industries. *Table 4* additionally shows the median daily wages of each nationality (in 1995 Euros). While Germans earn a median wage of 74 Euros in the *construction* sector, the corresponding values for foreigners are (i) generally significantly lower and (ii) still very heterogeneous. For instance, the median wage for Hungarians is 69 Euros, but for Czechs only 62 Euros. Of course, it can be assumed that such wage differentials are strongly affected by skill-specific differences or other differences in endowments. The following analyses take such differences into account.

[*Table 4* about here]

3 Econometric estimates

3.1 Outline of the estimation approaches

The descriptive evidence presented above reveals marked differences in the endowments of natives and immigrants. To explore the native-immigrant wage gaps further in a way which is consistent and meaningful in economic terms, we employ the decomposition method developed by Oaxaca (1973)

¹⁶ The corresponding share for Germans is 1.4 percent (the result is not included in *table 4*, but is available from the authors on request).

¹⁷ This sector is important for Slovaks, too. It may be argued that individuals working in *agriculture and hunting* are often employed as seasonal workers. We consider this point in a robustness check presented in the appendix.

and Blinder (1973)¹⁸. More specifically we ask whether lower wages are due to differences in characteristics or to other factors such as discrimination. Let the usual wage equation for the specific group of foreign workers and the reference group of German workers be given as $\ln w_i = x_i' \beta + \varepsilon_i$ and $\ln W_i = X_i' B + E_i$ respectively. Here, $\ln w_i$ (or $\ln W_i$ for Germans) stands for the logarithm of gross daily earnings¹⁹ for person i , and x_i (X_i) is a vector of individual and establishment level control variables. Specifically, we include qualification level, age, tenure (both with their squares), five establishment size categories, ten industry categories, year dummies and dummies controlling for the region and the region type.²⁰ Some complication in the decomposition is caused by a larger number of explanatory variables for immigrants. The regressions for foreigners additionally include entry year variables in order to account for potential differences in assimilation processes. This technical detail is not relevant, however, for understanding the results and is therefore shifted to the Appendix.

In the employment register, wages are censored at the upper earnings limit for social security contributions. To avoid bias and other complications when applying the decomposition method, wages are imputed in a preliminary step by estimating tobit regressions and replacing the censored wages with predictions from the tobit model. The dispersion of the wage distribution is preserved by adding random noise from a truncated normal distribution to the predicted values.²¹

To display the Oaxaca-Blinder decomposition concisely, we define $\Delta \hat{\beta} = \hat{\beta} - \hat{B}$ and $\Delta \bar{x} = \bar{x} - \bar{X}$, where the vectors \bar{x} and \bar{X} contain averages of the explanatory variables for immigrants and

¹⁸ Actually we apply the three-fold variant of the decomposition developed by Windsborough and Dickenson (1971). The label 'Oaxaca-Blinder decomposition' is, however, used instead since it is more established in the empirical literature.

¹⁹ Earnings are deflated to 1995 prices.

²⁰ We abstain from including a set of occupational indicators here. Besides the view that segregation into occupations can also be interpreted as discrimination (see e.g. Cain 1987), inclusion of both industry and occupation indicators is problematic because of high collinearity. This means that the two variables contain almost the same information, implying that the omission of occupations is harmless as long as we do not try to attribute wage effects to one of the two variables. All of the explanatory variables are described in *Table A1* in the appendix.

²¹ The statistical literature on multiple imputation shows that this procedure represents only residual uncertainty, neglecting coefficient uncertainty. Since the observation figures are quite large in our samples and the coefficients of the imputation model are estimated with considerable precision, coefficient uncertainty is negligible in our application and is therefore ignored.

Germans, respectively. Then the decomposition of the raw earnings differential $\ln \bar{w}_i - \ln \bar{W}_i$ has the form

$$\begin{aligned} \ln \bar{w}_i - \ln \bar{W}_i &= \bar{x} \hat{\beta} - \bar{X} \hat{B} \\ &= \underbrace{\bar{X} \cdot \Delta \hat{\beta}}_{\text{coefficients effect}} + \underbrace{\Delta \bar{x} \cdot \hat{B}}_{\text{characteristics effect}} + \underbrace{\Delta \bar{x} \cdot \Delta \hat{\beta}}_{\text{interaction effect}} \end{aligned} \quad (1)$$

The characteristics effect represents the log wage difference between immigrants and natives in a hypothetical situation where foreigners' skills are remunerated in the same way as those of natives (i.e. $\hat{\beta} = \hat{B}$). Analogously, the coefficients effect measures the log wage difference in a hypothetical situation where foreigners have the same characteristics on average as Germans. It is therefore sometimes interpreted as a direct measure of discrimination. This structural interpretation is only valid, however, if the regression model includes all relevant control variables. As unobserved heterogeneity remains in most practical applications, the label 'unexplained wage gap' is frequently employed in the literature instead. An alternative decomposition would replace the coefficients and mean characteristics of Germans with the corresponding foreigner values. This seems intuitively less meaningful in the case of immigrant wage gaps, however. Finally, the interaction effect measures the wage difference resulting for migrants if endowment differences were remunerated with coefficient differences. It is, however, of minor importance in our context and therefore not interpreted.

The mean decomposition results can only be taken as a representative description of the wage differences between immigrants and natives if the underlying data generating process is homoscedastic, i.e. if its coefficients do not vary across the wage distribution. Otherwise the means decomposition may conceal important differences across wage distributions. If, for example, discrimination has opposing signs at the upper and lower ends of the wage distribution (high-wage foreigners earn more and low-wage foreigners earn less than natives), the mean wage difference may be zero. To investigate the relevance of differential effects across the wage distribution, we repeat the decomposition using a quantile regression framework.²² Our implementation follows Melly (2005). First, we compute, for every country, 100 quantile regressions on an evenly spaced grid of quantiles.

²² For a similar application to U.S. data see Chiswick et al. (2008).

Though the censoring of wages could in principle be handled using the efficient three-step approach presented in Chernozhukov and Hong (2002), we avoid this additional computational burden by using imputed wages instead. This may introduce bias towards homoscedasticity as the imputation is based on standard (homoscedastic) tobit regressions. The bias remains negligible, however, if the decomposition results for quantiles close to the censoring limit are not interpreted. We choose quantile 0.8 as the upper limit for the results, which appears to be a safe limit for censoring shares below 10 percent. The quantile regression coefficients are then used (together with the characteristics) to construct counterfactual unconditional wage distributions. The quantile decomposition is analogous to the Oaxaca-Blinder decomposition, but counterfactual means are replaced by counterfactual unconditional distributions. The technical details of the decomposition are not required to understand the results. Interested readers are therefore referred to Melly (2005) or Angrist and Pischke (2009), section 7.2.

3.2 Decomposition results

Nationality-specific decomposition results for means

The estimation results of the specific wage equations differ for each nationality to some extent but are in line with the theoretical expectations and are therefore not presented here. *Table 5* contains the predicted log wage gaps between the specific group of immigrants and the reference group of Germans, as well as the results from the decomposition. It is obvious that the overall wage gaps differ significantly. The differential is largest for workers from Poland (-44 percent) and the Czech Republic (-38 percent). By contrast, Spaniards suffer only a moderate loss (-8 percent). At this stage of the analysis, however, it is not possible to conclude that workers from East European EU member states are generally worse off than workers from other areas. For instance, very large gaps are observed for workers from the EU member state of Portugal (-34 percent) and for Turks (-33 percent) and relatively small gaps for workers from the East European EU member countries Bulgaria (-17) and Hungary (-21 percent).

[*Table 5* about here]

As stated above, the overall wage gaps might be driven by the characteristics of workers. For instance, Spaniards are more highly qualified than Poles or Czechs, they work in different industries, regions and region types. Therefore, a more compelling approach takes into account the described differences and all other differences in observed characteristics. *Figure 1* depicts the decomposition results for each nationality.²³ Indeed, it can be observed that the characteristics effect (in absolute terms) is quite large for Poles (-23 percent) and small for Spanish immigrants (-4 percent). It can therefore be concluded that “poor” characteristics mainly contribute to explaining the overall wage differential for Poles. Besides Poland, there are three more countries (Portugal, Italy and Slovakia) where the characteristics effect dominates over the coefficients effect. For all other countries, the coefficients effect explains more than 50 percent of the overall wage gap. The unexplained gap (in absolute terms) is largest for immigrants from the Ukraine, Turkey (both -23 percent) and Belarus (-21 percent) and smallest for Bulgarians (-7 percent) and Spaniards (-9 percent). The coefficients effect is negative for all countries, however. The results strongly support the thesis that foreign workers are paid less than German workers even if they have observationally equivalent characteristics. If the influence of differences in unobserved characteristics is neglected,²⁴ our results suggest that immigrants are highly discriminated against in general. The nationality group “Other East” (Ukraine, Belarus and Russia) is affected most by discrimination. For the East European EU member states we observe discrimination effects which are comparable to those for the West and South European member states. These effects are also still smaller than for the last group, the Turks. In a more cautious interpretation of the coefficients effect, the conclusion would be reached that unobserved characteristics play an important role especially for the “Other East” immigrants and that they have a negative sign. This is plausible as this group appears to suffer from additional political and institutional restrictions, and a greater cultural distance (which may have direct wage effects in the labour market).

[*Figure 1* about here]

²³ For the sake of brevity, the interaction effect (presented in *table 5*) is not included in *figure 1*.

²⁴ From a theoretical point of view the influence of unobservable variables is ambiguous. On the one hand, variables which provide information on language skills or the transferability of human capital from the foreign labour market of origin into the German labour market could be expected to reduce the unexplained gap significantly. On the other hand, immigrants are assumed to be a highly-motivated sample of the foreign labour force. However, all studies on immigrant workers suffer from a lack of data in this respect.

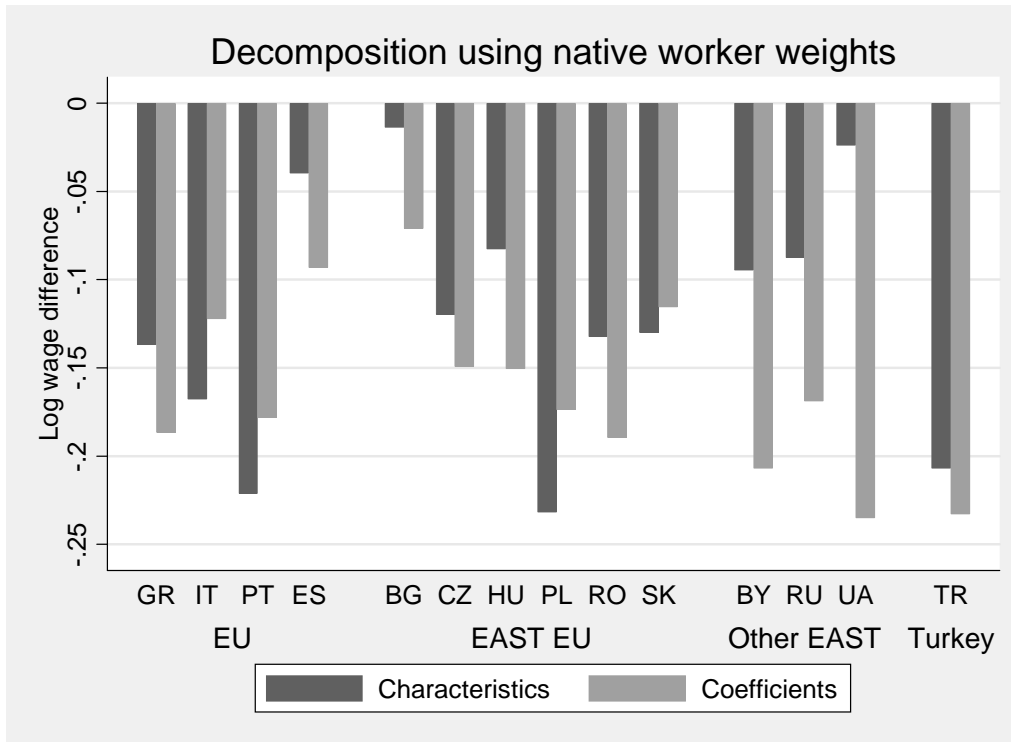


Figure 1: Nationality-Specific Decomposition Results (Using Native Worker Characteristics and Coefficients as Weights)

Before going on to interpret the results we pause for a moment to put our results in the context of the existing evidence. At first glance our wage differentials appear to be huge. On the basis of the GSOEP, Aldashev et al. (2008) find much smaller raw (log) wage differentials of 0.11 log points. A closer look at the samples suggests, however, that a great deal of the differences can be explained by differences in the samples. Firstly, Aldashev et al. (2008) have to include all entry cohorts due to observation number restrictions. We select immigrants entering Germany between 1995 and 2000 in order to restore comparability between West and East European nationalities at least to some degree.²⁵ Table 6 shows the importance of this. It contains decomposition results for the full samples 2001-2006 *without* entry year selection for nationalities with significant migrant inflows before the 1990s. A comparison with table 5 reveals considerably smaller wage gaps. The raw wage gaps decrease by about one third (in size) for all of the countries. Though it would go beyond the scope of this study to examine the assimilation effect for the entry cohorts in more detail, it is clear that immigrant wage gaps shrink as

²⁵ The immigrants from the former communist countries are pioneers in several respects whereas the immigrants from West European countries and Turkey join existing networks of compatriots with an established social infrastructure.

duration of residence in Germany increases (conditional on age and other controls). Secondly, the Aldashev et al. sample refers to all immigrants (including those from OECD countries) whereas we select specific nationalities for comparison purposes. Finally, the GSOEP is likely to suffer from survey sampling bias, which may be severe especially for immigrants, i.e. survey participation rates are likely to be higher for the more assimilated immigrants with above-average language skills. However, this group can be expected to earn higher wages, too. In summary, the comparison suggests firstly, that additional and in some cases more reliable demographic information from survey data (e.g. migration background, language skills, integration measures) is of limited value if used to explain a dependent variable where a large part of the total variation is removed or masked by survey sampling bias. And secondly, that register data play an important role in detecting sampling problems in survey data and in scaling up their results.

[Table 6 about here]

According to the immigration literature, the most reasonable explanation for the wage differential is that the international transferability of both formal education/training and labour market experience from the country of origin to the destination country is limited. Since formal education/training is less firm-, industry- and occupation-specific than labour market experience, the latter is assumed to be the main determinant. As a consequence, immigrants tend to be overeducated or overskilled for the jobs they do, i.e. they work in occupations whose skill requirements are lower than their own qualification level.²⁶ Systematically higher coefficients effects for Ukrainians, Belarusians and Russians suggest that the transferability is more problematic for workers from these countries than for others. A further explanation might be that the German language skills of these nationalities are less pronounced than in countries which are geographically closer to Germany.

Explanations of the immigrant wage gap which are based on search-theory models emphasise that immigrants who only stay in the host country for a limited number of years have less search experience and draw from a restricted sample of job offers. This can be viewed as a supportive rationale behind the huge wage gaps observed in our cohort. Furthermore it suggests additional wage losses for immigrants from East European countries since their residence permits are restricted by law.

²⁶ For more detailed explanations, see, for instance, Hartog (2000), Kiker et al. (2000) or McGuinness (2006).

This neither allows them to plan their careers in advance nor makes additional search efforts worthwhile. In this respect, the endowments of immigrants from Eastern countries are poorer both with regard to the past and their future prospects.

Industry-specific decomposition results for means

A glance at the industry affiliation of immigrants makes it clear that they are segregated into a handful of industries. To investigate whether sorting into industries may explain a large part of the coefficients effect, we run the decompositions on subsamples of workers in selected industries. The most important industries are *construction*, *hotels & restaurants* and *other business activities*. Because the *other business activities* sector covers very heterogeneous sub-sectors (see footnote 13), we concentrate on the first three sectors. Panel A of *table 7* contains the results for the *construction* sector. The number of observations is sufficiently large for Greeks, Italians, Portuguese, Czechs, Poles, Russians and Turks. The raw wage gap ranges from -13 percent (for Greeks, Portuguese and Poles) to -23 percent (Turks). This can be compared with the predicted differences for the seven countries in the aggregate, which are between -24 and -44 percent. The decomposition results also show that both the coefficients and the characteristics effects are distinctly smaller (in absolute terms) than in the aggregate. However, the coefficients effects dominate over the characteristics effects in each case. Furthermore, *relative* coefficients effects (computed as a fraction of the total wage differential) are even greater than in the aggregate. This indicates that discrimination plays a considerable role even within the construction sector, and that segregation or sorting into sectors contributes less to the explanation of immigrant wage differentials. Again, we find no evidence that East Europeans are more disadvantaged than other immigrants.

Turning to the *hotels & restaurants* sector (Panel B) the decomposition results corroborate the last findings. While the unexplained wage gap is -21 percent for Poles and -8 percent for Czechs, it is as large as -24 percent for Greeks and about -20 percent for Italians, Portuguese and Turks.

Nationality-specific decomposition results for quantiles

As mentioned above, mean decompositions only deliver representative and useful results if the true economic model is homoscedastic. To check this, we repeat the mean decomposition for the entire wage distribution using quantile regressions. As mentioned in the section describing the estimation approach above, the definition of characteristics and coefficients effects and the set of regressors are identical to those in the mean regression models. The only difference concerns sample size, and this difference is negligible. The sample size is cut here at 50,000 for countries exceeding this limit.²⁷ Thus the results can be compared directly with those from the mean regression models. To keep the exposition clear, only the results for the median and lower and upper quartiles are reported. Readers interested in the details may find the full decompositions presented as figures in the appendix. First we consider the predicted total differences. For example, the difference between the 25 percent quantiles of the Greek and the German (predicted) wage distributions is 30 log points, the corresponding difference for the 75 percent quantiles is 26 log points.

The differences between the immigrant and the German (log) wage distributions are larger at the lower quartile than at the upper quartile for most countries. This means that the (predicted) wage distributions of these countries are more widely dispersed than the German one. The differences between quartiles are most pronounced for Bulgaria (27 percentage points), Slovakia (13), Poland (12), Spain (12) and Hungary (10). Considerably less dispersed wage distributions are found especially for the Portuguese (-8), Belarusians (-5), Turks (-5) and Russians (-4). Next we inspect how the raw differences between quartiles translate into explained (characteristics) and unexplained (coefficients) parts. Regarding coefficients, we find pronounced decreasing effects (in absolute size) across the wage distribution for Bulgaria, the Czech Republic and Hungary. For example, the Czech coefficients effect decreases (in size) by 6 log points from -0.18 in the lower quartile to -0.12 in the upper quartile. If the estimation model is taken as a valid and sufficiently complete description of wages, the results imply that discrimination is more pronounced in the lower part of the distribution for these countries. An increase in discrimination across the wage distribution is interpreted in the literature as evidence of glass ceilings, a decrease as evidence of sticky floors (cf. Arulampalam et al.,

²⁷ The samples are of course drawn randomly.

2007). A thorough economic interpretation of these issues would require further examination of demographic characteristics for each nationality and a discussion of institutional aspects, however, which is clearly beyond the scope of this study.

Finally, regarding the characteristics effects, differences decrease considerably (in size) between quartiles for Spain (by 7 log points), Hungary (5 log points) and Slovakia (10 log points). This indicates that the characteristics of immigrants from these countries are more widely dispersed than those of native Germans. A considerable increasing characteristics effect (by 6 log points, from -0.16 for the lower to -0.22 for the upper quartile) can be found only for the Turk sample. The implied greater homogeneity of characteristics among Turks (especially regarding qualification level) is in concordance with our descriptive evidence.

In summary, the quantile decompositions add moderate qualifications to the mean decomposition results. Raw wage differences are more pronounced at lower quantiles for the majority of the countries considered, implying that the respective wage distributions are somewhat more widely dispersed than the German distribution. Similar patterns can be found for the coefficients and characteristics effects. But they are only pronounced for some of the countries.

4 Summary of findings

Particularly with regard to the full opening of the German labour market to East Europeans in 2011, this paper investigates whether East Europeans face particular disadvantages (in terms of wages) on the labour market. In order to detect systematic differences between nationalities, we analyse the immigrant wage gap for workers from the East European EU member countries and compare them with other nationalities. More specifically, we consider immigrants from “classical” EU member countries (Greece, Portugal, Italy and Spain), East European EU member states (Romania, Poland, Hungary, Czech Republic, Slovakia and Bulgaria), other East European countries (Ukraine, Belarus and Russia) and Turkey. Concentrating on a cohort of male workers who entered the German labour market between 1995 and 2000, we find that the overall wage gaps differ significantly. The differential

is largest for workers from Poland (-44 percent) and the Czech Republic (-38 percent) and by far the lowest for Spaniards (-8 percent).

Applying an Oaxaca-type decomposition technique, we split the earnings gap into an endowment effect and an unexplained effect. This provides evidence as to whether the negative differentials are due to unfavourable characteristics or remain unexplained by the model. The latter contains both discrimination and unobserved heterogeneity.. According to our results, unfavourable characteristics (compared with German workers) contribute significantly to explaining the immigrant wage gap. This is especially true for workers from Poland, Portugal, Italy and Slovakia. For all of the other countries the coefficients effect dominates. Even if attribution problems arise, the results suggest that immigrants are generally affected by discrimination. Contrasting the effects for workers from East European EU member countries with those for other nationality groups, it emerges that East Europeans are not worse off than others. The most pronounced indication for discrimination is found for immigrants from East European *non*-EU member states.

Taking a closer look at two sectors (*construction* and *hotels & restaurants*) in which foreigners typically work, we observe that the coefficients effect still contributes most to the explanation of the raw wage differential between foreigners and Germans. We therefore conclude that segregation into sectors does not significantly contribute to the discrimination of foreigners. The results rather suggest that foreigners are also affected by discrimination within sectors.

Results from quantile decompositions show that most immigrant wage distributions are somewhat more widely dispersed (i.e. differences from the native wage distribution are larger at lower quantiles), and that discrimination appears to be more pronounced at low wages for a majority of countries. Predicted differences (in absolute terms) decrease most across the wage distribution for Spain, Bulgaria, Poland and Slovakia. Coefficients effects (in absolute terms) decrease for several countries, especially for Bulgaria, the Czech Republic and Hungary. On the whole, our evidence suggests sticky floors rather than glass ceilings. The characteristics effects show a similar pattern for many countries,²⁸ indicating that endowments are more heterogeneous among immigrants than among Germans.

²⁸ A clear exception is only immigrants from Turkey, see above.

What can we conclude from our study regarding the opening of the German labour market to immigrants from the new EU member states in 2011? The focus of our paper does not permit comprehensive predictions about wage, distribution and welfare effects on the German labour market. Our results should rather be taken as a tessera. We find that immigrants from some East European countries show relatively good skill endowments. A comparison with other nationalities, however, shows no clear pattern which would make it possible to classify them as a homogenous group. The same result applies to the wages, i.e. characteristics and coefficients yield similar contributions for East European immigrants and other nationalities. Thus, immigration from the new EU countries will increase labour supply but the additional supply is characterized by similar ‘quality’ and can therefore be expected to yield similar wage effects for the immigrants.

References

- Angrist, J. and J. Pischke (2009), *Mostly Harmless Econometrics – An Empiricist’s Companion*, Princeton University Press, Princeton.
- Aldashev, A., J. Gernandt and S.L. Thomsen (2008), ‘The immigrant wage gap in Germany’, FEMM Working Paper Series 19, Faculty of Economics and Management, Magdeburg.
- Arulampalam, W., A.L. Booth and M.L. Bryan (2007), ‘Is there a glass ceiling over Europe? Exploring the gender pay gap across the wages distribution’, *Industrial and Labor Relations Review* 60, 163-186.
- Baas, T., H. Brücker and E. Hönekopp (2007), ‘Beachtliche Gewinne für die Deutsche Volkswirtschaft (Considerable Gains for the German Economy)’, IAB Kurzbericht Nr. 6/2007, IAB Nuremberg.
- Bartel, A. (1989), ‘Where do the new immigrants live?’ *Journal of Labor Economics* 7(4), 371-91.
- Bauer, T., G. Epstein and I. Gang (2005), ‘Enclaves, language, and the location choice of migrants’, *Journal of Population Economics* 18(4), 649-662.
- Bauer, T., G. Epstein and I. Gang (2002), ‘Herd Effects or Migration Networks? The Location Choice of Mexican Immigrants in the U.S.’, IZA Discussion Paper, 551.

- Bender, S., A. Haas and C. Klose (2000), 'The IAB employment subsample 1975-1995', *Schmollers Jahrbuch: Zeitschrift für Wirtschafts- und Sozialwissenschaften/Journal of Applied Social Science Studies*, 120(4), 649-662.
- Bender, S, J. Hilzendegen, G. Rohwer and H. Rudolph (1996), 'Die IAB-Beschäftigtenstichprobe 1975-1990: Eine Praktische Einführung', Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit, IAB, Nürnberg, BeitrAB 197.
- Blinder, A. (1973), 'Wage discrimination: reduced form and structural estimates', *Journal of Human Resources* 8, 436-455.
- Borjas, G.J. (1987), 'Self-Selection and the Earnings of Immigrants', *American Economic Review* 77(4), 531-553.
- Cain, G.G. (1987), 'The Economic Analysis of Labour Market Discrimination: A Survey', in: Ashenfelter, O., R. Layard and D. Card (eds.) *Handbook of Labour Economics*, Vol. 1. North-Holland: 693-781.
- Chernozhukov, V. and H. Hong (2002), 'Three-step censored quantile regression and extramarital affairs', *Journal of the American Statistical Association* 97, 872– 882.
- Chiswick, B.R. (1978), 'The Effect of Americanization on the Earnings of Foreign-Born Men', *Journal of Political Economy* 86(5), 81-87.
- Chiswick, B.R. and P.W. Miller (2007), 'The international transferability of immigrants' human capital', Discussion Paper 2670, IZA.
- Chiswick, B.R., A.T. Le and P.W. Miller (2008), 'How immigrants fare across the earnings distribution in Australia and the United States', *Industrial & Labor Relations Review* 61(3), 353-373.
- Diekmann, A., H. Engelhardt and P. Hartmann (1993), 'Einkommensungleichheit in der Bundesrepublik Deutschland: Diskriminierung von Frauen und Ausländern?', *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 26, 386-398.
- Gardeazabal, J. and A. Ugidos (2004), 'More on identification in detailed wage decompositions', *The Review of Economics and Statistics*, 86: 1034-1036.
- Haas, A. and J. Möller (2003), 'The agglomeration wage differential reconsidered - an investigation using German micro data 1983-1997'; in: Broecker, J., D. Dohse and R. Soltwedel (eds.), *Innovation clusters and interregional competition*, Berlin Heidelberg, New York, Springer.
- Hartog, J. (2000), 'Over-education and Earnings: Where Are We, Where Should We Go?' *Economics of Education Review* 19 (2), 131-147.
- Kiker, B.F., M.C. Santos and M. Mendes De Oliveira (2000), 'The Role of Human Capital and Technological Change in Overeducation', *Economics of Education Review* 19, 199-206.

- Lalonde, R.J. and R.H. Topel (1997), 'The Economic Impact of International Migration and the Economic Performance of Migrants', in: Rosenzweig, M.R. and O. Stark (eds.), *Handbook in Population and Family Economics*, Elsevier Science B.V.
- Lehmer, F. and J. Möller (2008), 'Interrelations between the urban wage premium and firm-size wage differentials- A micro data cohort analysis for Germany', *The Annals of Regional Science* (forthcoming).
- Melly, B. (2005), 'Decomposition of differences in distribution using quantile regression', *Labour Economics* 12, 577-590.
- McGuinness, S. (2006) 'Overeducation in the Labour Market', *Journal of Economic Surveys* 20 (3), 387-418.
- Nielsen, H.S., M. Rosholm, N. Smith and L. Husted (2004), 'Qualifications, discrimination, or assimilation? An extended framework for analysing immigrant wage gaps', *Empirical Economics* 29, 855-883.
- Oaxaca, R. (1973), 'Male-Female Wage Differentials in Urban Labor Markets', *International Economic Review* 14, 693-709.
- Windsborough, H.H. and P. Dickenson (1971), 'Components of Negro-White Income Differences', in: *Proceedings of the Social Statistics Section*, American Statistical Association, 6-8.
- Yun, M.-S. (2005), 'A Simple Solution to the Identification Problem in Detailed Wage Decompositions', *Economic Inquiry* 43, 766-772.
- Zavodny, M. (1999), 'Determinants of Recent Immigrants' Locational Choices', *International Migration Review* 33, 1014-1030.

Tables

Table 1: Observation Figures of Various Samples

Group	Country	A 95/06	A 01/06	E 01/06	E 01/06W	S 01/06W
EU	Germany	24,703,582	12,577,987	12,577,987	10,058,235	264,653*
	Greece	1,320,579	651,320	156,827	149,508	48,002
	Italy	2,462,941	1225,870	284,012	275,360	91,322
	Portugal	594,152	291,779	80,061	75,026	28,351
	Spain	507,658	239,291	44,853	42,883	11,496
EU East	Bulgaria	89,901	61,874	13,831	10,941	2,037
	Czech Rep.	144,991	89,681	48,062	45,571	17,531
	Hungary	161,167	76,651	18,090	15,735	4,163
	Poland	856,647	479,496	144,143	129,482	33,736
	Romania	273,382	144,329	44,922	42,916	11,010
	Slovakia	45,955	33,498	9,540	8,568	2,547
Other	Belarus	28,192	19,509	5,652	4,700	1,722
East	Russia	351,352	244,433	80,513	68,201	27,224
	Ukraine	156,459	119,265	39,387	32,631	10,914
Turkey	Turkey	6,869,725	3,477,755	922,848	875,418	264,653

Source: BEH, own calculations.

Legend: **A 95/06**: All observations 1995-2006. **A01/06**: All observations 2001-2006. **E 01/06** = A 01/06 restricted to individuals entering the German labour market between 1995 and 2000. **E 01/06W**: = E01/06 restricted to people in western Germany (who had never worked in eastern Germany). **S 01/06W**: = E 01/06W with additional restrictions (a) age between 25 and 55, (b) male, (c) in full-time employment, and (d) no apprenticeship training spell. **S 01/06W** is the final main estimation sample.

* To keep the estimation sample tractable, the German sample **S 01/06W** is restricted to the size of the largest group of foreigners (Turks).

Table 2:
Nationality-Specific Distribution of Gender (Entry-Sample 1995/2000)

Group	Country	Male	Female	Total
EU	Germany	5,217,555	4,840,680	10,058,235
		52%	48%	100%
	Greece	79,638	69,870	149,508
		53%	47%	100%
	Italy	154,645	120,715	275,360
	56%	44%	100%	
	Portugal	40,123	34,903	75,026
		53%	47%	100%
	Spain	18,905	23,978	42,883
		44%	56%	100%
EU East	Bulgaria	3,630	7,311	10,941
		33%	67%	100%
	Czech Rep.	22,819	22,752	45,571
		50%	50%	100%
	Hungary	5,875	9,860	15,735
		37%	63%	100%
	Poland	45,526	83,956	129,482
35%		65%	100%	
Romania	14,570	28,346	42,916	
	34%	66%	100%	
Slovakia	3,457	5,111	8,568	
	40%	60%	100%	
Other East	Belarus	2,301	2,399	4,700
		49%	51%	100%
	Russia	35,397	32,804	68,201
		52%	48%	100%
	Ukraine	16,438	16,193	32,631
50%		50%	100%	
Turkey	Turkey	482,486	392,932	875,418
		55%	45%	100%
Total		6,143,365	5,691,810	11,835,175

Source: BEH, own calculations.

Note: The comparison group of German workers is a randomly drawn sample whose size is equal to the sum of all immigrant worker samples considered in the respective years. The observation figures are given in column **E 01/06W** of *table 1*.

Table 3:
Nationality-Specific Distributions of Skill Level (Estimation Sample)

Group	Country	Missing	Low-skilled	Skilled	Highly skilled	Total
EU	Germany	9.8%	10.5%	67.0%	12.7%	264,653
	Greece	24.6%	39.1%	29.4%	6.9%	48,002
	Italy	25.5%	33.0%	36.1%	5.4%	91,322
	Portugal	31.7%	36.8%	29.4%	2.2%	28,351
	Spain	18.2%	17.8%	41.4%	22.6%	11,496
EU East	Bulgaria	22.4%	16.8%	31.6%	29.1%	2,037
	Czech Rep.	15.5%	21.2%	58.4%	5.0%	17,531
	Hungary	22.4%	14.4%	47.6%	15.6%	4,163
	Poland	31.5%	29.1%	34.8%	4.5%	33,736
	Romania	22.6%	27.9%	41.6%	7.9%	11,010
	Slovakia	28.5%	16.7%	40.1%	14.8%	2,547
Other East	Belarus	25.6%	24.2%	34.3%	16.0%	1,722
	Russia	22.5%	28.9%	33.7%	15.0%	27,224
	Ukraine	19.7%	19.3%	40.3%	20.7%	10,914
Turkey	Turkey	26.6%	39.3%	32.1%	2.0%	264,653

Source: BEH, own calculations.

Note: Observation figures are given in column **S 01/06W** of *table 1*.

Table 4:
Nationality-Specific Differences with Respect to Industry Affiliation (Estimation Sample)

Group	Country	Top 1 Industry	(P50)	Share 1	Top 2 Industry	(P50)	Share 2	Top 3 Industry	(P50)	Share 3
EU	Germany	45 Construction	(73.72)	8.7%	74 Oth. Business Act.	(80.96)	7.4%	29 Mach. Equipment	(93.07)	6.8%
	Greece	55 Hotels Restaurants	(31.14)	11.2%	74 Oth. Business Act.	(52.43)	10.6%	34 Motor Vehicles	(91.35)	7.5%
	Italy	55 Hotels Restaurants	(36.63)	16.5%	45 Construction	(66.84)	11.4%	74 Oth. Business Act.	(53.15)	7.3%
	Portugal	45 Construction	(65.97)	26.1%	74 Oth. Business Act.	(51.46)	11.2%	55 Hotels Restaurants	(38.56)	10.2%
	Spain	74 Oth. Business Act.	(70.42)	11.9%	51 Wholesale Trade	(72.41)	7.3%	55 Hotels Restaurants	(41.01)	5.8%
EU East	Bulgaria	55 Hotels Restaurants	(45.10)	18.3%	74 Oth. Business Act.	(86.47)	11.2%	92 Recreation	(74.52)	7.9%
	Czech Republic	45 Construction	(62.34)	10.9%	28 Fabr. Metal Prod.	(52.69)	10.6%	15 Food Beverages	(48.06)	9.7%
	Hungary	55 Hotels Restaurants	(45.93)	11.5%	74 Oth. Business Act.	(64.24)	9.8%	45 Construction	(68.67)	9.7%
	Poland	1 Agriculture Hunting	(38.79)	22.5%	45 Construction	(66.25)	10.1%	92 Recreation	(32.62)	5.9%
	Romania	45 Construction	(66.09)	8.2%	55 Hotels Restaurants	(42.94)	7.5%	74 Oth. Business Act.	(58.23)	7.1%
	Slovakia	55 Hotels Restaurants	(46.64)	24.0%	1 Agriculture Hunting	(53.19)	8.2%	51 Wholesale Trade	(78.02)	5.5%
Other East	Belarus	45 Construction	(64.49)	12.7%	74 Oth. Business Act.	(59.67)	8.2%	28 Fabr. Metal Prod	(66.84)	7.0%
	Russia	74 Oth. Business Act.	(47.38)	11.3%	45 Construction	(63.17)	7.4%	28 Fabr. Metal Prod	(64.68)	6.8%
	Ukraine	74 Oth. Business Act.	(54.99)	14.2%	85 Health	(75.24)	6.5%	51 Wholesale Trade	(65.58)	6.5%
Turkey	Turkey	74 Oth. Business Act.	(48.56)	12.6%	45 Construction	(62.11)	8.0%	34 Motor Vehicles	(89.39)	7.2%

Source: BEH, own calculations.

Notes: Observation figures are given in column **S01/06 W** of *table 1*. P50 (in parentheses) denotes median daily wages (in 1995 Euros) for each nationality in the specific sector.

Legend: Const.: Construction; Oth. Business Act.: other business activities; Hotels Restaurants: hotels & restaurants; Mach. Equipment: manufacture of machinery & equipment; Health: health & social work; Fabr. Metal Prod.: manufacture of fabricated metal products; Food Beverages: manufacture of food products and beverages; Recreation: recreational, cultural and sporting activities; Motor Vehicles: manufacture of motor vehicles

Table 5: Decomposition of Log Real Wages at Sample Means (Comparison Group: Germans)

Group	Country	Pred. Difference	Coeff. Effects	S.D.	Chara. Effects	S.D.	Interaction Effects
EU	Greece	-0.29	-0.19	0.003	-0.14	0.001	0.03
	Italy	-0.29	-0.12	0.002	-0.17	0.001	0.00
	Portugal	-0.34	-0.18	0.005	-0.22	0.002	0.06
	Spain	-0.08	-0.09	0.005	-0.04	0.003	0.05
EU East	Bulgaria	-0.17	-0.07	0.012	-0.01	0.006	-0.09
	Czech Rep.	-0.38	-0.15	0.005	-0.12	0.002	-0.11
	Hungary	-0.21	-0.15	0.010	-0.08	0.004	0.02
	Poland	-0.44	-0.17	0.003	-0.23	0.002	-0.03
	Romania	-0.30	-0.19	0.007	-0.13	0.002	0.02
	Slovakia	-0.26	-0.12	0.010	-0.13	0.006	-0.02
Other East	Belarus	-0.21	-0.21	0.025	-0.09	0.006	0.09
	Russia	-0.24	-0.17	0.007	-0.09	0.002	0.02
	Ukraine	-0.22	-0.23	0.017	-0.02	0.003	0.04
Turkey	Turkey	-0.33	-0.23	0.002	-0.21	0.001	0.11

Source: BEH, own calculations.

Notes: All comparisons relate to the German sample. **Pred. Diff.** is the predicted mean (log) wage difference obtained from the regression model. **Coeff. Effects** denotes the predicted counterfactual mean (log) wage difference that would prevail if the characteristics of foreigners were identical to those of German workers. **Chara. Effects** denotes the predicted counterfactual mean (log) wage difference that would prevail if the coefficients of foreigners were identical to those of German workers. **S.D.** denotes standard errors obtained from 100 bootstrap replications.

Observation figures are given in column **S01/06W** of *table 1*. All estimation results relate to the sample of male immigrants in full-time employment who entered the western German labour market between 1995 and 2000. The estimation period is 2001-2006.

Table 6. Decomposition of Wage Differences for the Sample Including All Entry Cohorts (EU Countries and Turkey Only)

Group	Country	Pred. Difference	Coeff. Effects	S.D.	Chara. Effects	S.D.	Interaction
EU	Greece	-0.20	-0.16	0.001	-0.08	0.001	0.04
	Italy	-0.19	-0.11	0.001	-0.10	0.000	0.01
	Portugal	-0.22	-0.14	0.001	-0.14	0.001	0.06
	Spain	-0.06	-0.05	0.001	-0.03	0.001	0.02
Turkey	Turkey	-0.21	-0.18	0.001	-0.12	0.000	0.09

Source: BEH, own calculations.

Notes: All comparisons relate to the German sample. All estimation results relate to the sample of male immigrants in full-time employment. The estimation period is 2001-2006.

Legend: **Pred. Diff.** is the predicted mean (log) wage difference obtained from the regression model. **Coeff. Effects** denotes the predicted counterfactual mean (log) wage difference that would prevail if the characteristics of foreigners were identical to those of German workers. **Chara. Effects** denotes the predicted counterfactual mean (log) wage difference that would prevail if the coefficients of foreigners were identical to those of German workers. **S.D.** denotes standard errors obtained from 100 bootstrap replications.

Table 7: Decomposition of Log Real Wages (Relative to Germans) in the Sectors Construction and Hotels & Restaurants

Group	Country	Pred. Difference	Coeff. Effects	Chara. Effects	Interaction
<i>Panel A: Construction</i>					
EU	Greece	-0.19	-0.11	-0.08	0.00
	Italy	-0.13	-0.07	-0.07	0.01
	Portugal	-0.13	-0.07	-0.06	-0.01
	Spain	-	-	-	-
EU East	Bulgaria	-	-	-	-
	Czech Rep.	-0.22	-0.11	-0.02	-0.09
	Hungary	-	-	-	-
	Poland	-0.13	-0.08	-0.06	0.00
	Romania	-	-	-	-
	Slovakia	-	-	-	-
Other East	Belarus	-	-	-	-
	Russia	-0.17	-0.11	-0.06	0.00
	Ukraine	-	-	-	-
Turkey	Turkey	-0.23	-0.13	-0.11	0.01
<i>Panel B: Hotels & Restaurants</i>					
EU	Greece	-0.44	-0.24	-0.19	-0.02
	Italy	-0.33	-0.20	-0.21	0.08
	Portugal	-0.29	-0.20	-0.16	0.07
	Spain	-	-	-	-
EU East	Bulgaria	-	-	-	-
	Czech Rep.	-0.17	-0.08	-0.05	-0.04
	Hungary	-	-	-	-
	Poland	-0.30	-0.21	-0.14	0.06
	Romania	-	-	-	-
	Slovakia	-	-	-	-
Other East	Belarus	-	-	-	-
	Russia	-	-	-	-
	Ukraine	-	-	-	-
Turkey	Turkey	-0.27	-0.21	-0.14	0.08

Source: BEH, own calculations.

Notes: All comparisons relate to the German sample. **Pred. Diff.** is the predicted mean (log) wage difference obtained from the regression model. **Coeff. Effects** denotes the predicted counterfactual mean (log) wage difference that would prevail if the characteristics of foreigners were identical to those of German workers. **Chara. Effects** denotes the predicted counterfactual mean (log) wage difference that would prevail if the coefficients of foreigners were identical to those of German workers. **S.D.** denotes standard errors obtained from 100 bootstrap replications.

Again, the base sample is **S01/06W** of *table 1*. Results are reported only if the number of nationality-specific observations in a given sector is larger than 1,000.

Table 8: Quantile Decomposition of Log Real Wages (Relative to Germans)

Group	Country	PD25	PD50	PD75	CO25	CO50	CO75	CH25	CH50	CH75
EU	Greece	-0.30	-0.25	-0.26	-0.21	-0.17	-0.18	-0.12	-0.12	-0.13
	Italy	-0.30	-0.25	-0.25	-0.15	-0.12	-0.11	-0.15	-0.15	-0.16
	Portugal	-0.30	-0.32	-0.38	-0.19	-0.18	-0.19	-0.19	-0.20	-0.23
	Spain	-0.13	-0.09	-0.01	-0.11	-0.09	-0.07	-0.07	-0.05	0.00
EU East	Bulgaria	-0.29	-0.17	-0.02	-0.11	-0.07	-0.06	-0.07	-0.01	0.06
	Czech Rep.	-0.38	-0.41	-0.42	-0.18	-0.15	-0.12	-0.11	-0.11	-0.13
	Hungary	-0.27	-0.22	-0.17	-0.17	-0.15	-0.13	-0.10	-0.08	-0.05
	Poland	-0.50	-0.39	-0.38	-0.20	-0.17	-0.18	-0.23	-0.21	-0.21
	Romania	-0.28	-0.27	-0.31	-0.18	-0.19	-0.22	-0.12	-0.12	-0.12
	Slovakia	-0.34	-0.27	-0.21	-0.14	-0.11	-0.11	-0.17	-0.13	-0.07
Other East	Belarus	-0.21	-0.25	-0.26	-0.22	-0.25	-0.28	-0.11	-0.10	-0.07
	Russia	-0.24	-0.26	-0.27	-0.18	-0.20	-0.21	-0.10	-0.09	-0.06
	Ukraine	-0.23	-0.23	-0.20	-0.23	-0.24	-0.26	-0.05	-0.02	0.01
Turkey	Turkey	-0.30	-0.30	-0.34	-0.25	-0.23	-0.25	-0.16	-0.18	-0.22

Source: BEH, own calculations.

Notes: All comparisons relate to the German sample. **PDQ** is the predicted (log) wage difference at quartile Q (Q=25,50,75) obtained from the regression model. **COQ** denotes the predicted counterfactual (log) wage difference at quartile Q that would prevail if the characteristics of foreigners were identical to those of German workers. **CHQ** denotes the predicted counterfactual (log) wage difference at quartile Q that would prevail if the coefficients of foreigners were identical to those of German workers.

The observation figures are the same as in table 2 if they do not exceed 50,000. Otherwise 50,000 observations are drawn randomly from the nationality sample concerned. All estimation results relate to the sample of male immigrants in full-time employment who entered the western German labour market between 1995 and 2000. The estimation period is 2001-2006.

Appendix

**Table A1:
A Description of the Variables**

<i>Variable</i>	<i>Description</i>
log wage	Logarithm of gross daily earnings, calculated as the average over the observed employment period for each person (in 1995 prices).
qualification	The qualification level is divided into four categories: (i) low-skilled (reference group): people with no occupational qualification regardless of which schooling level, that means with or without upper secondary education (Abitur); (ii) skilled: people with an occupational qualification regardless of whether they have completed upper secondary education (Abitur); (iii) highly-skilled: people with upper secondary education and a degree from a university or university of applied science; (iv) skill missing: persons with missing information on skills.
age	Linear and quadratic terms of age; measured in years.
tenure	Linear and quadratic terms of tenure; measured in years.
establishment size	Establishment size is divided into five categories. Using the cumulative distribution of establishment size, the categories are generated to enclose an identical number of observations. The categories are: (i) 0 – 20 percent (reference category) (ii) 20.01 - 40 percent (iii) 40.01 - 60 percent (iv) 60.01 - 80 percent (v) 80.01 - 100 percent.
industry	The industry classification (based on WZ 03) comprises ten broad categories.
region	We consider five regional units: (i) North (reference category): Schleswig-Holstein, Hamburg, Bremen, Lower-Saxony; (ii) NRW: North Rhine-Westphalia; (iii) Centre: Rhineland-Palatinate, Hesse, Saarland; (iv) BW: Baden-Württemberg; (v) Bavaria.
region type	For our estimates we use a classification scheme developed by the Federal Office for Building and Regional Planning (<i>Bundesamt für Bauwesen und Raumordnung</i> - BBR) which comprises nine different types: (i) metropolitan core cities (reference category), (ii) highly urbanized, (iii) urbanized and (iv) rural districts in areas with large agglomerations; (v) central cities, (vi) urbanized and (vii) rural districts in regions with features of conurbation; (viii) urbanized and (ix) rural districts in regions of rural character.
entry year	Six dummies capture the exact year of entry into the German labour market in the period 1995 (reference year) to 2000 (included for immigrant samples only).
time dummies	Six year dummies capture the exact time of observation in the period 2001 (reference year) to 2006.

Treatment of additional regressors in the immigrant samples

Our models contain entry year dummies for immigrants only. This generates a minor technical problem for the decompositions since the characteristics and coefficients effects are based on differences between *all* regressors and coefficients. Dropping the entry year dummies before performing the decomposition would lead to biased results since the constant of the immigrant sample

is shifted by these dummies. The constant coefficient of the foreigner sample represents the intercept of the basis category and thus cannot be compared with the native sample coefficient representing the intercept for the whole native sample. This problem can be avoided by applying the restriction that the weighted sums of coefficients of each dummy variable set equals zero.²⁹ Then the dummy coefficients can be interpreted as deviations from a common mean and the constant coefficient represents the intercept of the whole sample. Restricted estimation was proposed by Gardeazabal and Ugidos (2004) and Yun (2005) to solve the more general problem of the decomposition results being affected by the choice of the base category in the presence of dummy regressors (even for equal numbers of regressors in both samples). Thus we apply the restrictions to *all* dummy variable sets. In the conditional mean models this is performed using restricted least squares. Since our implementation of quantile regressions does not permit the inclusion of the restrictions directly, we instead reparametrize the dummy variables in the following way. Consider the regression equation

$$w = x\beta + \sum_{k=1}^K D_k \gamma_k + \varepsilon$$

where $\sum_{k=1}^K D_k \gamma_k$ contains all the dummies of a dummy variable group (e.g. entry year) and $x\beta$ contains all other regressors (including the constant) and ε is the residual. The

restriction $\sum_{k=1}^K w_k \gamma_k = 0$ can be solved for the last dummy coefficient $\gamma_K = -\frac{1}{w_K} \sum_{k=1}^{K-1} w_k \gamma_k$.

Back substitution into $\sum_{k=1}^K D_k \gamma_k$ gives

$$\sum_{k=1}^K D_k \gamma_k = \sum_{k=1}^{K-1} D_k \gamma_k - \frac{1}{w_K} \sum_{k=1}^{K-1} w_k \gamma_k = \sum_{k=1}^{K-1} \left(D_k - \frac{w_k}{w_K} D_K \right) \gamma_k = \sum_{k=1}^{K-1} \delta_k \gamma_k.$$

Thus the restricted regression is performed by substituting the original dummies D_k by the

reparametrized ones $\delta_k = D_k - \frac{w_k}{w_K} D_K$, $k = 1, \dots, K-1$, in the regression equation. And the

²⁹ The weights are computed as shares of the respective dummies in the total sample.

coefficient of the omitted $K - th$ dummy is obtained from the restriction. The reparametrized entry year dummies are included in the regressions but do not enter the decomposition.

Robustness checks:

1. Nationality coding problems

Nationality is reported less reliably than other variables in the employment register data since it is not required for determining pension entitlements and unemployment benefits. A thorough examination of the nationality variable reveals that there are individuals who are registered as foreigners in some years and as Germans in others. These changes often occur after establishment switches and show no clear pattern. Most importantly, as returns to the former foreign nationality are frequent, they do not indicate that the individuals concerned obtained German citizenship. A plausible explanation for these German spells is that employers tick German nationality (the default) for foreigners if it is difficult or inconvenient to find out the correct nationality. Note that our results are severely biased if Germans are wrongly coded as foreigners whereas the reverse case (foreigners coded as Germans) is almost harmless since the number of foreigners is small (compared with the number of Germans). Consequently foreigners added (incorrectly) to the German sample would only have a minor impact on the German sample coefficients and mean characteristics.

To check the impact of nationality coding errors, we construct a worst-case sample where all foreigners with at least one spell of German nationality in the estimation period 2001-2006 and foreigners with at least two apprenticeship spells in Germany are dropped completely from the estimation sample.³⁰ The following table shows the observation figures and decomposition results for this ‘worst case’ sample. Comparison with the main results³¹ in table A2 shows only minor differences, supporting the hypothesis that the spells with German nationality are mainly coding errors.

³⁰ Apprenticeship spells in Germany indicate foreigners who grew up in Germany and may thus be very similar to natives with respect to qualification level and language skills. Note that this problem is more severe for the EU countries and Turkey.

³¹ Note that the immigrant subsamples of the main estimation sample do not contain German spells. For example if a person has five Czech spells and one German spell in the period 2001-2006, only the five Czech spells are used in the main decomposition.

Table A2: Mean Decomposition Results from the Worst Case Sample

Column No.	(1)	(2)	(3)	(4)	Estimation Sample (5)	Comparison Samples (6)	(7)
Group	Country	Pred. Diff.	Coeff. Effects	Chara. Effects	Only For. No App.	Only For.	S 01/06 W
EU	Greece	-0.35	-0.20	-0.16	26,187	31,022	48,002
	Italy	-0.36	-0.16	-0.19	42,259	50,416	91,322
	Portugal	-0.36	-0.20	-0.23	18,203	20,088	28,351
	Spain	-0.03	-0.11	0.03	5,684	7,545	11,496
EAST EU	Bulgaria	-0.21	-0.10	-0.02	1,262	1,284	2,037
	Czech Rep.	-0.40	-0.14	-0.14	11,960	12,039	17,531
	Hungary	-0.22	-0.14	-0.08	2,688	2,766	4,163
	Poland	-0.54	-0.20	-0.27	19,537	20,346	33,736
	Romania	-0.36	-0.22	-0.15	5,991	6,174	11,010
	Slovakia	-0.29	-0.13	-0.15	1,799	1,823	2,547
Other EAST	Belarus	-0.18	-0.26	-0.06	852	885	1,722
	Russia	-0.22	-0.16	-0.06	14,062	14,357	27,224
	Ukraine	-0.24	-0.34	-0.02	6,128	6,358	10,914
Turkey	Turkey	-0.37	-0.26	-0.22	146,181	185,764	264,653

Notes: Column (7): S01/06W contains the observation figures of the standard estimation sample for the decomposition results in table 5 (means) and table 7 (quantiles) above. Column (6): ‘Only For.’ is obtained from sample S01/06W by dropping all people with at least one German spell in 2001-2006.

The observation figures in column (5): ‘Only For., No App.’ are obtained from column (6) by dropping all people with at least two apprenticeship spells in Germany. Column (5) represents the observation figures relevant for the decomposition results in columns (2)-(4). Decomposition results for sample (6) are similar and are therefore not reported here but are available on request from the authors.

Legend: **Pred. Diff.** is the predicted mean (log) wage difference obtained from the regression model. **Coeff. Effects** denotes the predicted counterfactual mean (log) wage difference that would prevail if the characteristics of foreigners were identical to those of German workers. **Chara. Effects** denotes the predicted counterfactual mean (log) wage difference that would prevail if the coefficients of foreigners were identical to those of German workers.

2. Qualification coding problems

As mentioned in the data description, the reliability of the qualification variable is poorer for immigrant workers for several reasons. Firstly, they are more likely to accept jobs below their (formal) qualification level to compensate for other disadvantages or because their outside options are poorer (e.g. entitlements to unemployment benefits). Though the reporting procedure asks for the employees’ highest formal qualification level, employers appear sometimes to report the qualification level required to perform the job under consideration. Together, employment below the actual qualification level and this reporting bias generate a negative qualification bias. Secondly, grades and occupational qualifications gained in foreign countries are frequently not recognized or approved in Germany. Employers can be expected to report lower qualification levels or missing information if the approval

of qualifications from the countries of origin is unclear. Thirdly, the qualifications of immigrants may differ from those of native workers even if they have identical formal grades and the qualifications are reported correctly. Even then, the training in the country of origin may have been conducted using different technologies or may have focused on different aspects than in Germany. Since measures to check or improve the reliability of the qualification variable are limited, we follow an indirect strategy to obtain information about the size and the sign of the implied bias by repeating the decomposition for a subsample where reporting of the qualification variable can be expected to be more reliable than average. The factors mentioned above suggest that the categories ‘low-skilled’ and ‘missing’ are chosen by employers when the qualification level is unclear. Thus the information appears to be most reliable where workers are classified as medium or highly skilled. However, as censoring is more severe for the highly skilled and overqualification is more likely for them, we drop them too. This leaves us with a sample restricted to medium skilled workers (those who have completed an apprenticeship). If the reported qualification level systematically understates immigrants’ actual level, the decomposition delivers a larger characteristics effect and a smaller coefficients effect (both in absolute values). Consequently the (relative) importance should decrease for the characteristics and increase for the coefficients effects in the restricted sample. The comparison corroborates this, i.e. the relative³² characteristics effects are about 15 percentage points smaller and the relative coefficients effects are about 20 percentage points larger in the restricted sample.³³ Thus, though the qualification coding problem is likely to generate estimation bias, the bias is moderate. More importantly, the coefficients effects are understated and would even increase in the absence of coding errors.

3. Seasonal work effects

As documented above (see *table 4*), the most important sector for Polish workers is *agriculture / hunting*. It can be observed that this is also the second most important sector for Slovaks. As seasonal work plays a central role in this sector, it may be argued that the results for these two nationalities are not comparable with those for other nationalities. Unfortunately, information on seasonal work is not

³² Relative effects are computed as shares of the predicted differences.

³³ The detailed results are not reported to save space but are available from the corresponding author upon request.

reported directly in the data source. We therefore resort to durations of the employment spells as a proxy. It emerges, in fact, that the employment spells of Poles and Slovaks are distinctly shorter on average than those of other nationalities. Consequently, as robustness check we restrict the analysis to employment spells lasting longer than 90 days and repeat the decomposition of sample means. Compared with the results documented above, both the predicted differences for Poles (-0.37 instead of -0.44) and Slovaks (-0.23 instead of -0.26) as well as the characteristics effects (-0.20 instead of -0.23 for Poles; -0.11 instead of -0.13 for Slovaks) decrease to some extent. The coefficients effect and all effects for other nationalities remain unaffected. After dropping *additionally* the sector *agriculture / hunting* from the sample, the predicted differences for Poles decrease to -0.32 and the characteristics and coefficients effects to -0.15 and -0.16, respectively. The corresponding measures for Slovaks are -0.21, -0.10 and -0.12. In summary, the robustness check adds a qualification to the results for Poles. Besides that, it supports our view that notwithstanding the uncertainty regarding seasonal work in our data, the representativeness of the documented results is unharmed.

Figure A1: Quantile Decomposition Results for all Countries

