THE UNIVERSITY OF MELBOURNE

Melbourne Institute Working Paper Series

Working Paper No. 14/08

Labour Market Outcomes of Second Generation Immigrants: How Heterogeneous Are They Really?

Stefanie Schurer

MELBOURNE INSTITUTE
of Applied Economic and Social Research

Labour Market Outcomes of Second Generation Immigrants: How Heterogeneous Are They Really?*

Stefanie Schurer
Melbourne Institute of Applied Economic and Social Research
The University of Melbourne

Melbourne Institute Working Paper No. 14/08

ISSN 1328-4991 (Print)

ISSN 1447-5863 (Online)

ISBN 978-0-7340-3282-9

August 2008

* This study was funded by the Krupp von Bohlen und Halbach Foundation as part of a PhD degree at the Ruhr Graduate School in Economics, Essen, Germany. Many thanks go to Thomas Bauer, Jan Brenner, John Creedy, Annika Frohloff, Paul H. Jensen, Wang-Sheng Lee, Mathias Sinning, Christoph M. Schmidt, Elizabeth Webster and participants of the Microeconometrics Seminar at the University of Melbourne for valuable comments.

Melbourne Institute of Applied Economic and Social Research
The University of Melbourne
Victoria 3010 Australia
Telephone **(03) 8344 2100**
Fax **(03) 8344 2111**
Email **melb-inst@unimelb.edu.au**
WWW Address **http://www.melbourneinstitute.com**

Abstract

This study comprehensively portrays the labour market outcomes of second generation immigrants in Germany. Special attention is attributed to *observable* heterogeneity in terms of country of origin and *unobservable* heterogeneity in terms of parental human capital, neighbourhood effects, and mixed marriage background. Pooled, static and dynamic panel data models, and a decomposition analysis are used to estimate and explain the average differences in hourly wages and unemployment probabilities separately for men and women. The results suggest that the second generation cannot be considered as one homogeneous group: some groups perform better, equally or worse than comparable German natives. Also, relative outcomes in wages depend mainly on observable characteristics, whereas relative unemployment risks are mainly driven by unobservable factors.

In Germany, second generation immigrants are making up a sizeable fraction of the overall population, nevertheless very little is known about their labour market outcomes. Existing evidence of poor schooling outcomes suggests that second generation immigrants in Germany will also perform poorly in local labour markets. We know from the empirical literature that second generation immigrants in the United States are perceived to perform better, whereas in European countries they are perceived to perform worse than their native peers. Much of this can be traced to the relatively high levels of education among second generation immigrants in the United States. These differences suggest that investment into human capital and the transition from school to work differ between second generation immigrants and comparable natives (Card and Schmidt, 2003). From a theoretical perspective, however, it is surprising that education levels and subsequent labour market success should systematically differ, as children of immigrants have the chance to undergo the same educational institutions as native children. This should be particularly the case for Germany, where education is free.

This paper thus explores whether second generation immigrants in Germany follow the European trend of falling behind the labour market success of their native peer-group. Wage earnings and unemployment risks of second generation immigrants are compared to those of German natives using data from the German Socio-Economic Panel (GSOEP). A second generation immigrant is identified on the basis of non-German nationality and arriving in Germany no later than six years of age. The main assumption made to conduct the analysis is that labour market outcomes of second generation immigrants are heterogeneous. A large part of this heterogeneity is hypothesised to be captured by the country or origin of the parents[1]. Country of origin proxies differences in skill endowment and attitude to acquire locally valued human capital. The other part of heterogeneity in labour market outcomes is assumed to be attributed to differences in the intergenerational transmission of productivity, parental human capital, neighbourhood effects, and mixed marriage backgrounds. These individual-specific factors are mainly unobserved in the GSOEP, and therefore they may confound statistical inference of the effect of migration background on labour market outcomes.

To control for individual-specific heterogeneity, I augment simple pooled OLS and Pro-

[1]Schmidt (1992) concludes that labour market outcomes of first generation immigrants in Germany differ across countries of origin.

bit models with random effects Mundlak specifications that allow for potential correlation between the individual specific effects and regressors of the model (Chamberlain, 1980; Mundlak, 1978). Some of these unobserved factors, e.g. parental human capital and the intergenerational transmission of productivity, may be the cause of the autoregressive nature of wages and the state dependence of unemployment. Thus, I re-estimate the random effects Mundlak model with a linear dynamic specification for wages and a Wooldridge (2005) conditional maximum likelihood approach for unemployment probabilities. Further, to understand the main contributors to wage and unemployment differences between German natives and second generation immigrants, a threefold decomposition analysis is conducted.

The estimation results suggest that the second generation cannot be considered as one homogeneous group. Some groups perform better in terms of hourly wages, while other groups perform worse in finding employment than German natives. Unobserved heterogeneity plays a strong role in determining unemployment risk differentials, but the main story about wage differentials is told by observable characteristics.

The remainder of the paper is structured as follows: in the next Section the main results from the empirical literature are reviewed. Section 2 outlines a simple theoretic framework to sort out the mechanisms underlying the relationship between country of origin, parental human capital, neighbourhood effects, and the degree of migration background and labour market outcomes. Section 3 explains the econometric specification for both hourly wages and unemployment risks and the methods chosen to control for unobserved heterogeneity and dynamics in the data. Section 4 introduces the data set and the variables used for the analysis. In Section 5 a descriptive analysis of the data and the main results of the estimated models of the hourly wages and unemployment risks are presented, which also includes a decomposition analysis of outcome differentials. Section 6 tests for the robustness of the estimated labour market advantages and disadvantages of some groups. Section 7 concludes.

1 Empirical Evidence

Second-generation immigrants, the children of original immigrants, are constituting or becoming to constitute a sizeable fraction of the North American and younger European

population. There are good reasons to understand the economic performance of second generation immigrants: their educational attainment and economic performance are an important indicator for how successful a host country is in integrating its immigrant population in the long run.

The main conclusion from the North American literature is that second generation immigrants are perceived as relatively successful in the US and Canada due to their high levels of education (Aydemir and Sweetman, 2006; Card, 2005). In contrast, second generation immigrants in European immigration countries seem to fall behind the labour market success of their native peer-groups, but performance varies across countries of origin (Behtoui, 2004; Jakobsen and Smith, 2003; Nielsen et al., 2003; Van Ours and Veenman, 2003). An overarching theme of the literature is that the intergenerational transmission of labour market outcomes, human capital attainment, parental education, and neighbourhood effects are crucial determinants of the labour market position of immigrant children (Card and Schmidt, 2003).

In Germany, second generation immigrants are estimated to make up 6 % of the total population (OECD, 2005). A recent study by the OECD (2006) concludes that Germany is one of the countries in which 15-year old second generation immigrants achieve lower schooling results than the first generation and far lower results than the German native average. Out of the 17 countries considered by the study, Germany performs worst in integrating second generation immigrants. The OECD's claim, that immigrant children fall behind their native counterparts in terms of educational achievement, has been shown by several other German studies. Their main message is that children with migration background have a lower probability of obtaining a higher secondary school qualification[2] and an unambiguously higher probability of obtaining only the minimum schooling requirement (Gang and Zimmermann, 2000; Haisken-DeNew et al., 1997; Kristen, 2000; Riphahn, 2003). This trend seems to widen over time (Riphahn, 2005).

According to human capital theory (Becker, 1964), education levels are crucial in determining labour market outcomes. If second generation immigrants lag behind their German native counterparts in terms of educational outcomes, they are also likely to lag

[2]In Germany three different secondary high school qualifications are available. The minimum schooling qualification is nine years of schooling (Hauptschule). There is an intermediate schooling qualification that requires ten years of schooling and enables its graduates to pursue further education excluding the access to university. The highest secondary schooling qualification requires 13 years of schooling and enables its graduates to enter university.

behind in their labour market outcomes.

For Germany, there is little empirical evidence on the labour market outcomes of second generation immigrants, though. Fertig and Schmidt (2001) provide a portrait of first and second generation immigrants using the 1995 wave of the Mikrozensus. The empirical analysis focuses mainly on the welfare dependence of immigrants. For second generation immigrants, they observe a pattern of welfare dependence which is similar to that observed among native Germans. A recent OECD study finds that the higher probability of being unemployed among second generation immigrants can be fully explained by their low educational qualifications (OECD, 2005). Uhlendorff and Zimmermann (2006) report that second generation Turkish immigrants face a higher probability of unemployment than their comparable German counterparts. A descriptive analysis of a unique data set collected in Nuremberg in 1999, the EFFNATIS field study, reports that it is mainly second generation Turkish immigrants who cluster in low skilled and routine positions and who face the highest unemployment rates (Worbs, 2003). Gestring et al. (2004), who interviewed 55 Turkish second generation immigrants in a medium-sized town in Lower Saxony, suggest that Turkish offspring are poorly integrated. The authors find that Turkish second generation immigrants end up in low and unskilled work, enter the job market without any qualifying degree, engage in discontinuous or temporary employment, and experience long periods of unemployment.

All of these studies use relatively small samples or look only at one part of labour market outcomes (e.g. only unemployment). It remains an open question whether these results are representative for the average second generation immigrant.

2 Theoretical Framework

An individual's wage depends on his or her productivity P and on business fluctuations in the economy B:

$$W = f(P, B), \tag{1}$$

where $f(.)$ is an undefined function with regular properties. While B is assumed to affect natives (from here onwards referred to as N) and second generation immigrants (from here onwards referred to as SGI) in the same way, I assume that productivity P is determined differently for the two groups. The productivity for natives P^N is a function of education

(E), experience (EXP), and labour market relevant individual skills (U) such as cognitive ability or motivation:

$$P^N = f(E, EXP, U). \tag{2}$$

Haveman and Wolfe (1995) and Becker and Tomes (1979) suggest that the most fundamental factor in describing a child's educational attainment is the parent's human capital, and that there is a strong correlation between parents' and children's lifetime earnings and wealth (Behrman and Taubman, 1976). For this reason, life-time achievements of the parents and its transmission to their children need to be accounted for.

In this paper, the intergenerational transmission of productivity is understood as the influence of the father's life-time productivity, proxied by lifetime wages, on the *dependence of current productivity on past productivity* of the native:

$$P_t^N - \sigma P_{t-1}^N = f(\frac{1}{T_f} \sum P_t^{Father}, X), \tag{3}$$

where $0 \leq \sigma \leq 1$. σ is the strength of the dependence of productivity between periods and therefore determines implicitly the productivity growth rate. P_t^{Father} is the father's productivity in any time-period t, T_f is the number of years the father worked, $f(.)$ is an unspecified function, and X are other determinants. Since the individual's productivity growth rate depends on the father's lifetime productivity (among others), the individual's current productivity (wage) depends on both father's lifetime productivity $(\frac{1}{T_f} \sum P_t^{Father})$ and the past period productivity (P_{t-1}^N).

Current period productivity of natives increases with education, experience, ability and motivation, a low dependence parameter, and a low intergenerational transmission of productivity.

For SGI, the productivity and intergenerational transmission equations are similar to Eqs. (2) and (3), except for their augmentation by a cultural difference parameter K[3]:

$$P^{SGI} = f(E, EXP, U, K). \tag{4}$$

K reflects the fact that SGI experience a different preparation to enter the labour market than German natives despite facing the same educational opportunities. The choice of

[3]This cultural difference parameter has been proposed by Bratsberg et al. (2006) in the context of wage differentials of immigrants to the United States.

education may also be a function of the cultural difference indicator, but to keep the analysis straight-forward, I assume that education is exogenous.

Cultural differences are a function of country of origin differences (C), parental human capital (PHC), neighbourhood effects (NHE), and the degree of migration background (M):

$$K = f(C, PHC, NHE, M). \tag{5}$$

Country of origin of the parents is a good summary statistics for the cultural difference, as it provides insight about the linguistic difference between the mother tongue of the parents and the native language, the religion and values SGI most likely will acquire. These are factors which influence the integration process. Country of origin differences are also a good proxy for the average skill level of the parents, as country of origin is an indicator for the reason of migration. For instance, during the guest-worker period mainly low skilled workers were attracted to come to Germany and they were recruited from Southern European countries.

The lower the parental human capital and the larger the proportion of immigrants in the neighbourhood in which the SGI grew up, the larger cultural differences are expected to be. Parental human capital, e.g. language skills and knowledge about education opportunities, plays an important role in facilitating assimilation of SGI in schools and development of language ability. Neighbourhood effects, i.e. the average socioeconomic achievement of the surrounding environment or ethnic peer group, affect the potential to acquire local labour market skills (Borjas, 1992, 1993). A similar effect can be expected from the degree of migration background of SGI. Children from mixed marriage backgrounds, i.e. one parent is a native, are more likely to acquire local labour market skills than SGI whose parents are both foreigners[4].

The productivity of a SGI is, ceteris paribus, a negative function of the cultural difference parameter K: the greater the difference in cultural background to the host country's required level of local labour market skills, the lower the expected productivity level of the SGI:

[4]Behtoui (2004) has shown that immigrant children from families in which one parent, mainly the father, was a native showed less labour market disadvantages than children from pure immigrant backgrounds.

$$\frac{\partial P^{SGI}}{\partial K} < 0. \tag{6}$$

The resulting wage differentials between natives and SGI are the differences in productivity:

$$\begin{aligned} W^N - W^{SGI} &= f(P^N, B) - f(P^{SGI}, B), & (7)\\ &= f(E, EXP, U; B) - f(E, EXP, U, K; B). & (8) \end{aligned}$$

For convenience, I assume that education, experience and unobserved factors motivation and ability U are exogenous and the same for both natives and SGI, then we get:

$$W^N - W^{SGI} = f(K) = f(C, PHC, NHE, M). \tag{9}$$

Eq. (9) states that, ceteris paribus, the expected wage differentials between natives and SGI are the greater the larger the country of origins difference, the lower the parental human capital of the SGI, the greater the density of immigrants in the neighbourhood in which the SGI grew up, and the stronger the migration background.

The differences in wages are a positive function of the cultural difference parameter K, as:

$$\frac{\partial (W^N - W^{SGI})}{\partial K} = -\frac{\partial P^{SGI}}{\partial K} > 0. \tag{10}$$

The last inequality follows from Eq. (6). Similar arguments can be made for the differences in unemployment risks between German natives and SGI.

3 Empirical Specification

3.1 Earnings

The empirical model builds on the standard human capital earnings function of Becker and Chiswick (1966) and Mincer (1974). Suppose the wage equation of an individual observed in calendar year t is:

$$y_{it} = \alpha + \sum_g D_{ig}\beta_g + AG'_{it}\delta + X'_{it}\theta + \Pi'_{it}\mu + u_{it}. \tag{11}$$

where y_{it} is the logarithm of real hourly wage of person $i = 1, \ldots, N$ in year $t = 1, \ldots, T_i$ (unbalanced panel); X is a vector of socioeconomic characteristics such as schooling, vocational training, marital status, household composition, and the state in which the person resides at time t. AG takes the value 1 if the person belongs to the age-group in calendar year t. Age-group indicators are preferred over linear or quadratic age profiles to capture nonlinearities in wage growth, particularly observed for women. Π denotes a set of indicator variables set to unity if the observation is made in calendar year t. The group of second generation immigrants is captured by the indicator D_{ig}, which takes the value 1 if the individual belongs to a group g stemming from a certain country (or group of contries) and 0 otherwise. How many elements are in g depends on the degree of heterogeneity, observable in terms of country of origin, assumed for that group. From here onwards the status of being a second generation immigrant is interpreted as *migration background*. The error term u_{it} captures all unobserved factors that are assumed to vary independently of the regressors of the model.

In the empirical analyses, I estimate wage equations separately by gender. In a first step, Eq. (11) is estimated with pooled Ordinary Least Squares (POLS) excluding all variables but the second generation sub-group dummies. Subsequently, regressors are added to the base model, exploring which observable characteristics explain wage differences between natives and SGI. According to the theoretical framework, the main parameter vector of interest is β, as it captured observable cultural differences in productivity.

According to Eqs. (2) and (4), productivity depends also on individual-specific motivation or ability. Productivity of SGI further also depends on parental human capital, neighbourhood effects, and whether they stem from a mixed marriage background, as assumed in Eq. (5). These factors are unobservable in my data-set. To control for their presence and possible correlation with the regressors of the model, I allow α in Eq. (11) to vary across individuals and allowing α_i, the time-invariant unobserved heterogeneity, to correlate linearly with the mean values of the time-varying regressors of the model. Chamberlain (1980) and Mundlak (1978) show that, in the case of linear models, fixed effects and random effects are numerically identical if the correlation between α_i and all right-hand-side variables W_{it} take the following linear form:

$$\alpha_i = \bar{W}_i \zeta + r_i, \tag{12}$$

r_i is a random effect such that $E[r_i|\bar{W}_i] = E[r_i] = 0 \ \forall \ t$, $\bar{W}_i = \frac{1}{T_i}\sum_{i=1}^{T_i} W_{it}$, and T_i denotes the number of observations of respondent i in the sample. In order to identify β_g, the impact of the time-invariant migration background, I additionally assume that $E[r_i|W_{it}, D_{ig}] = E[r_i] = 0 \ \forall \ t$. This assumption states that the unobserved time-invariant effect does not correlate systematically with the country of origin of the individual. For the case of ability, this assumption explicitly rules out politically controversial conjectures such as *immigrants from certain countries are less able or less motivated.* Thus, ability or motivation vary systematically across individuals but not across country of origin.

This approach allows for correlation between α_i and all other regressors in the model while being able to identify the effect of country of origin on labour market outcomes[5]. Replacing α_i with Eq. (12) yields for each time period t:

$$y_{it} = \sum_g D_{ig}\beta_g + AG'_{it}\delta + X'_{it}\theta + \Pi'_{it}\mu + (\bar{W}_i\zeta + r_i) + u_{it}. \tag{13}$$

In a third step, the term $\psi_1 y_{it-1}$, i.e. the lagged value of the hourly wage, is added to the random effects specification. This dynamic specification captures the autoregressive nature of labour market processes.

$$y_{it} = \sum_g D_{ig}\beta_g + AG'_{it}\delta + X'_{it}\theta + \Pi'_{it}\mu + \psi_1 y_{it-1} + (\bar{W}_i\zeta + r_i) + u_{it}. \tag{14}$$

The dynamic approach is justified on the ground that lagged wages capture an additional source of unobserved heterogeneity that conventional random effects Mundlak specifications cannot pick up. As theoretically argued in Eq. (3), state dependence of earnings is influenced among others by the intergenerational transmission of productivity.

Dynamics in the wage determination process are also emphasised in the wage curve debate. In a review of the literature, Blanchflower and Oswald (2005) conclude that the coefficient on the lagged value of wages is approximately 0.5. Baltagi et al. (2007) find that wages exhibit a high degree of autoregression both at the regional and individual level in Germany. Preferring a dynamic over a static specification of wages using the IAB Employment Panel, they estimate a coefficient on the lagged variable in the order of 0.5.

[5]Allowing for this correlation in a more flexible specification would require fixed effects estimation, but both dummy variable least square and within estimators do not allow identification of β_g.

3.2 Unemployment

For the unemployment equation, a model similar to Eq. (11) is specified. Let UE^*_{it} be the true, but unobserved, individual propensity of unemployment. This latent propensity is assumed to be a linear function of observable characteristics and an error term:

$$UE^*_{it} = \alpha + \sum_g D_{ig}\beta_g + AG'_{it}\delta + X'_{it}\theta + \Pi'_{it}\mu + u_{it}, \tag{15}$$

in which the variables and vectors are defined as in Eq. (11). The latent propensity is not directly observable, but the indicator of being unemployed UE_{it} is observable. It takes the value 1 if the true underlying propensity of unemployment is greater than a certain threshold level which is normalised to 0, and 0 otherwise:

$$UE_{it} = \begin{cases} 1 & \text{if } UE^*_{it} > 0 \\ 0 & \text{if } UE^*_{it} \leq 0. \end{cases}$$

Assuming the error term to be normally distributed ($u_{it} \sim N(0,1)$) yields the probability of being unemployed:

$$Pr(UE_{it} = 1) = \Phi(\alpha + \sum_g D_{ig}\beta_g + AG'_{it}\delta + X'_{it}\theta + \Pi'_{it}\mu + u_{it}), \tag{16}$$

where Φ denotes the cumulative standard normal distribution function. Parameter estimates are obtained by Maximum Likelihood.

Analogous to the estimation strategy for the linear wage regression, regressors are added to the base model to test whether the raw differentials in unemployment risks for various migration backgrounds disappear once comparing the comparable. In the first step, the model is estimated by a pooled Probit approach implicitly assuming no intercept heterogeneity. In the second step, time-invariant unobserved heterogeneity is assumed and controlled for by re-estimating the unrestricted model in Eq. (16) with a random effects specification that includes a Mundlak specification of the error term (same arguments as in the wage equation). When summarising all regressors of the model with Z_{it}, the average values of the time-variant regressors with $\bar{W}_i$, and all parameter vectors with ξ,

the random effects probit model can be written as:

$$f(UE_1 \dots UE_T|Z_i, \xi) = \int_{-\infty}^{+\infty} \Big\{ \prod_{t=1}^{T} f(UE_{it}|Z_{it}, \bar{W}_i, r_i, \xi) \Big\} \frac{1}{\sigma_r} \phi(\frac{r}{\sigma_r}) dr, \tag{17}$$

by assuming $r_i|Z_i, \bar{W}_i \sim N(0, \sigma_r^2)$, $u_{it} \sim N(0,1)$, and $f(UE_{it}|Z_{it}, \bar{W}_i, r_i, \xi) = \Phi(\alpha + \sum_g D_{ig}\beta_g + AG'_{it}\delta + X'_{it}\theta + \Pi'_{it}\mu + (u_{it} + \bar{W}_i\zeta + r_i))^{UE_{it}}[1 - \Phi(\alpha + \sum_g D_{ig}\beta_g + AG'_{it}\delta + X'_{it}\theta + \Pi'_{it}\mu + (u_{it} + \bar{W}_i\zeta + r_i))]^{1-UE_{it}}$. Taking the logarithm of Eq. (17) gives the conditional log likelihood $L_i(\xi)$ for each individual i. The log-likelihood function for the entire sample N can be maximized with respect to ξ and σ_r^2 to obtain the $\sqrt{N}$-consistent asymptotically normal estimator. The relative importance of the unobserved effect in the total variation of the model is measured as $\rho = \frac{\sigma_r^2}{1+\sigma_r^2}$, which can also be interpreted as the correlation the composite latent error $(u_{it} + \bar{W}_i\zeta + r_i)$ bewteen any two time periods (Wooldridge, 2002, p. 486).

In the third step, the state dependence of unemployment is controlled for. Empirically, it has been observed that experiencing unemployment in one year makes future unemployment more likely (Clark and Summers, 1979; Layard et al., 1991). Sweeney (1996) found for British data that about half of those leaving unemployment have a high probability of relapsing into unemployment within a year.

Random effects models with dynamics introduce yet another source of bias due to the presence of the time-invariant unobserved effect. If the unobserved heterogeneity exhibits persistence over time, then ignoring it will lead to an overstatement of the true persistence in unemployment. To estimate the model, an assumption is required about the initial observation UE_{i1} and in particular about its relationship with the unobserved heterogeneity component. I follow Wooldridge (2005)'s Conditional Maximum Likelihood (CML) approach that considers the distribution conditional on the initial period value of unemployment and exogenous variables. Instead of modelling $f(UE_{i1}, \dots, UE_{iT}|Z_i, \bar{W}_i)$, Wooldridge suggests modelling $f(UE_{i2}, \dots, UE_{iT}|UE_{i1}, Z_i, \bar{W}_i)$. This produces a simple estimation method which has the advantage that it can be implemented with standard random effects probit software.

To make the marginal effects of the random effects probit comparable to those of the pooled probit, the coefficients have to be re-scaled by $\sqrt{1-\hat{\rho}}$ (Arulampalan, 1999), where $\hat{\rho} = \frac{\hat{\sigma}_r^2}{1+\hat{\sigma}_r^2}$.

3.3 Decomposition of Outcome Differences

To understand whether the differences in the hourly wages and unemployment risks between second generation immigrants and German natives are mainly driven by differences in the observable characteristics or by the differences in unobservable characteristics, I apply the decomposition method proposed by Oaxaca (1973) and Blinder (1973). In particular, I use the extension by Daymont and Andrisani (1984) that allows decomposing the outcome differentials of interest not only into differences in endowment and coefficients, but also in differences in interactions between endowments and coefficients. One advantage of this approach is to interpret the outcome differentials exclusively from the perspective of one of the two groups of interest.

The decomposition is calculated from the perspective of the second generation immigrant (SGI), i.e. choosing the coefficient β_{SGI} and the values of the observable characteristics X^{SGI} as benchmark:

$$y^N - y^{SGI} = (X^N - X^{SGI})\beta_{SGI} + X^{SGI}(\beta_N - \beta_{SGI}) + (X^{SGI} - X^N)(\beta_{SGI} - \beta_N),$$

and changing the sign of the third term on the right-hand side, one gets:

$$y^N - y^{SGI} = (X^N - X^{SGI})\beta_{SGI} + X^{SGI}(\beta_N - \beta_{SGI}) - (X^N - X^{SGI})(\beta_N - \beta_{SGI}). \quad (18)$$

In Eq. (18) $y^N - y^{SGI}$ is the difference in predicted outcomes between German natives (N) and second generation immigrants (SGI). The first term on the right-hand side $(X^N - X^{SGI})\beta_{SGI}$ is interpreted as how would the predicted outcome of a second generation immigrant have changed if he or she had the same endowments as a German native. The second term $X^{SGI}(\beta_N - \beta_{SGI})$ says how the predicted outcome y^{SGI} would have changed if the second generation immigrant had the same unobserved characteristics, implied by the difference in coefficients, as a German native. The last term on the right-hand side $(X^N - X^{SGI})(\beta_N - \beta_{SGI})$ states how the predicted outcome of the second generation immigrant would have changed if he or she had the same observable and unobservable characteristics as the German native.

My analysis for the unemployment probabilities is conducted with an implementation of the generalised decomposition method for non-linear models in STATA as described by

Bauer and Sinning (2008b,a)[6].

4 Data

The sample for the empirical analysis comes from 22 waves (1984-2005) of the German Socio-Economic Panel (GSOEP), which is a representative survey conducted annually[7].

Two sets of restrictions are used to include an individual into the sample. First, a second generation immigrant is identified as an individual born in Germany after 1954[8] and having a foreign nationality[9] (from here onwards this is referred to as the strict definition). Alternatively, the second generation is identified as an individual born abroad after 1949, who enters Germany no later than six years of age (from here onwards this is referred to as the wide definition). The latter definition is widely used in the German literature, e.g. Worbs (2003). It assumes that pre-school cultural knowledge is not indicative for assimilation behaviour. Second, German natives are selected into the sample only, if they were born in the same birth-year interval as all second generation immigrants, if they were born in Germany, and if they have German citizenship. Individuals are selected from 16 years onwards.

The following classification of having a migration background is used. Assuming complete homogeneity across country of origin, *all second generation immigrants* are classified within one group. Assuming moderate heterogeneity, one may distinguish between *Turks*, *Guest-workers excluding Turks*, and *Non guest-workers*. If one assumes a greater extent of observable heterogeneity, country of origin may be further distinguished between *Turks*, *Yugoslavs*, *Greeks*, *Italians*, *Spanish or Portuguese*, *EU 15 members*, and *Non EU 15*

[6]The analysis for both the linear and nonlinear decomposition is conducted with the *nldecompose.ado* in STATA. Many thanks to the authors Thomas Bauer, Markus Hahn, and Mathias Sinning of the RWI Essen for providing their code. The same results for the linear decomposition are obtained by using the *oaxaca* command written by Ben Jann (Jann, 2008).

[7]The data used in this paper was extracted from the SOEP Database provided by the DIW Berlin (http://www.diw.de/soep) using the Add-On package PanelWhiz v1.0 (Oct 2006) for Stata(R). PanelWhiz was written by Dr. John P. Haisken-DeNew (john@panelwhiz.eu). The PanelWhiz generated DO file to retrieve the SOEP data used here and any Panelwhiz Plugins are available upon request. Any data or computational errors in this paper are our own. Haisken-DeNew and Hahn (2006) describe PanelWhiz in detail.

[8]This date has been chosen to ensure that the guest-worker children are identified correctly. The recruitment program started in 1955 with Italy.

[9]We identified an individual's nationality by the nationality he or she held in the first interview. Immigrant children who changed nationality are separately considered in a robustness check. Another solution would have been to use parental birthplace or nationality. This information is, unfortunately not available.

members. Which stratification is most appropriate is an empirical matter.

Various indicators for wages are tested to compute the hourly wage, but I use the logarithm of the ratio of *individual labour earnings*, which consists of the gross annual salary of the main job and *yearly hours worked*. This variable reflects the sum of all monthly salaries before tax deduction. It comprises bonus payments such as holiday bonus, and the so-called 13^{th} and 14^{th} monthly salary and thus captures times of unemployment or underemployment.

To compute the indicator of unemployment I use *being registered as unemployed*, as this measure provides insight into the direct costs of unemployment to society. To be able to receive unemployment benefits an individual needs to register with a local unemployment agency.

Age-group indicators take the value 1 if the individual is within a specific age-group and 0 otherwise. The brackets for the age-groups run from 21 to 25, 26 to 30, 31 to 35, 36 to 40, 41 to 50, and 51 to 60. The base category is 16 to 20 years of age, a category in which more than 32 % of the second generation is classified. Five-year intervals at the younger age-groups are utilised because slightly more than 80 % of the second generation immigrants are no older than 30 years.

Socioeconomic characteristics are captured by dummy variables for education and vocational training. Individuals who have no more than ten years of schooling and who did not engage in vocational training are the base category. The first dummy variable includes those individuals who hold the basic or the intermediary school qualification from *Hauptschule* or *Realschule*, respectively, and who completed an apprenticeship. The second dummy variable represents those individuals who have finished the university qualifying secondary degree *Abitur* but who did not attend university. The third dummy variable represents all individuals who have obtained a university degree.

Parental background information is captured in indicator variables that proxy the socioeconomic status of the father (mother) when the individual was 15 years of age and the religion of the father (mother). The former is the variable *job position of father at the age of 15*. The indicators are *Father was a blue collar worker*, *Father was a white collar worker* (which also includes civil servants), *Father was inactive*, *Father was self-employed*, and *No answer on father's job position*. The base category is *Father was a blue collar worker*. Another measure available is parental schooling. However, the information in the

GSOEP on parents' education is noisy because a large proportion of immigrants' degrees are classified as *other degree*, which may be any degree between primary and tertiary school level. Where the individual is living is proxied by the state dummy variables (*Bundesländer*). These state dummy variables and time dummy variables capture business cycles and local labour market conditions.

I do not include self-assessed language proficiency as an explanatory variable proposed by Dustmann (1994) and applied by Constant and Massey (2005). These subjective measures of language proficiency are prone to misclassification error and thus estimated coefficients may be severely biased (Dustmann and Van Soest, 2001). Moreover, language proficiency may be endogenous with respect to labour market earnings.

5 Results

5.1 Descriptive Analysis

One of the main problems in analysing the labour market outcomes of second generation immigrants is that the sample sizes available in the data are small. Table 1 reports the number of individuals available in the sample for each sub-group of second generation immigrants for both the strict and the wide definition. The strict definition refers to all second generation immigrants who were born in Germany, whereas the wide definition considers all immigrants who arrive at an age no older than 6 in Germany. The wide definition of second generation yields a sample of roughly 1,266 second generation immigrants, 447 Turks, 195 Yugoslav, 185 Greek, 247 Italian, 134 Spanish or Portuguese, and 58 from non guest-worker backgrounds. Given that the strict definition yields a sample that is only two thirds of that size, I continue to work with the wide definition from here onwards. From the small sample sizes of the non guest-worker groups (58 individuals), it is self-explanatory that a further distinction between EU and non EU backgrounds is statistically not feasible.

Table 1: Number of individual for each group in sample

	SG	Turk	Yugo	Greek	Ital	Span	GW	NGW	EU	NONEU
Strict: born in Germany	874	285	143	136	182	92	518	36	27	9
Wide: arrived in Germany at age < 6	1,266	447	195	185	247	134	723	58	40	18

Tables 2 and 3 show the number of person-year observations that have a positive

outcome for four different employment status measures for German natives and second generation immigrants for men and women, respectively. The numbers in Column 1 make it clear that for Spanish men and women and individuals belonging to the group of non guest-workers the outcome measure of *being registered unemployed* counts a small number of *positive outcomes* (e.g. 23 (22) Spanish men (women) report being registered unemployed, and 32 (9) non guest-worker men (women) report being registered unemployed). For all other sub-groups the number of *positive outcomes* is reasonably large of a minimum of 50 observations.

Using the outcome measure currently not working yields samples large enough for differentiating the analysis between Turkish, Italian, Spanish, Greek, Yugoslav, and non guest-worker second generation immigrants. This measure is used in the sensitivity analysis.

Table 2: Sample size information for male individuals by employment status

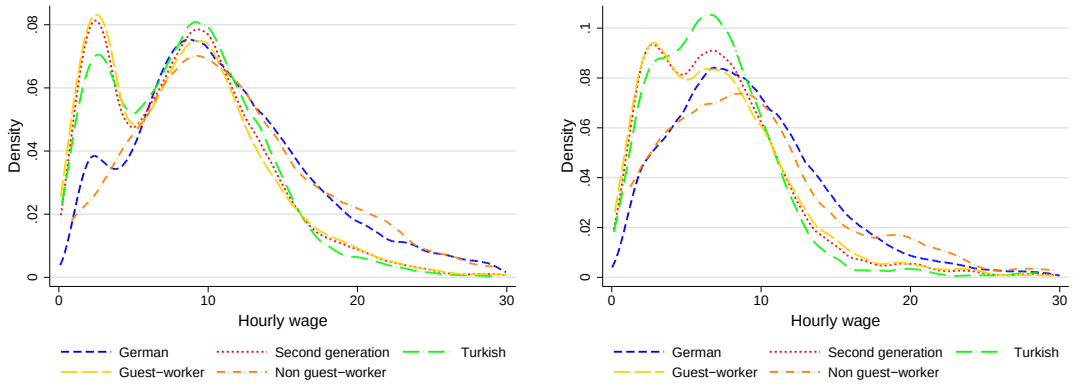
	No wages observed		Positive wages	
	Registered unemployed	Currently not working	Employed full-time	part-time
German	4,613	11,699	53,475	6,671
Second generation	468	1,499	2,933	976
Turkish	259	677	1,006	338
Guest-worker	167	713	1,702	567
Yugoslav	56	236	348	178
Greek	46	206	409	124
Italian	52	185	635	204
Spanish	23	121	361	97
Non guest-worker	32	74	174	35
European Union	29	59	127	21
Non European Union	3	15	47	14

Table 3: Sample size information for female individuals by employment status

	No wages observed		Positive wages	
	Registered unemployed	Currently working	Employed full-time	part-time
German	5,576	23,312	28,450	22,594
Second generation	323	2,066	1,499	1,040
Turkish	131	893	481	295
Guest-worker	183	1,071	946	678
Yugoslav	66	272	266	190
Greek	43	312	224	169
Italian	53	382	362	255
Spanish	21	132	105	77
Non guest-worker	9	75	61	54
European Union	5	48	47	42
Non European Union	4	27	14	12

Fig. 1 depicts the distribution of hourly wages for men and women. For both men and

women, the distribution of hourly wages is similar between German natives and second generation immigrants from non guest-worker countries. For these two groups the income distribution is slightly shifted to the right, indicating a greater proportion of individuals with a wage beyond the most common value. For the second generation as a homogeneous group, and equally for Turkish and guest-worker men, the distribution of wages is bimodal. One part of the second generation obtains wages similar to the most common hourly wage of German natives. Another part obtains wages significantly below the German natives' modal value. In particular, the distribution of wages of Turkish men and women is slightly shifted to the left. These observations from the raw data suggest a first guess that a large proportion of second generation immigrants cluster in low income sectors.

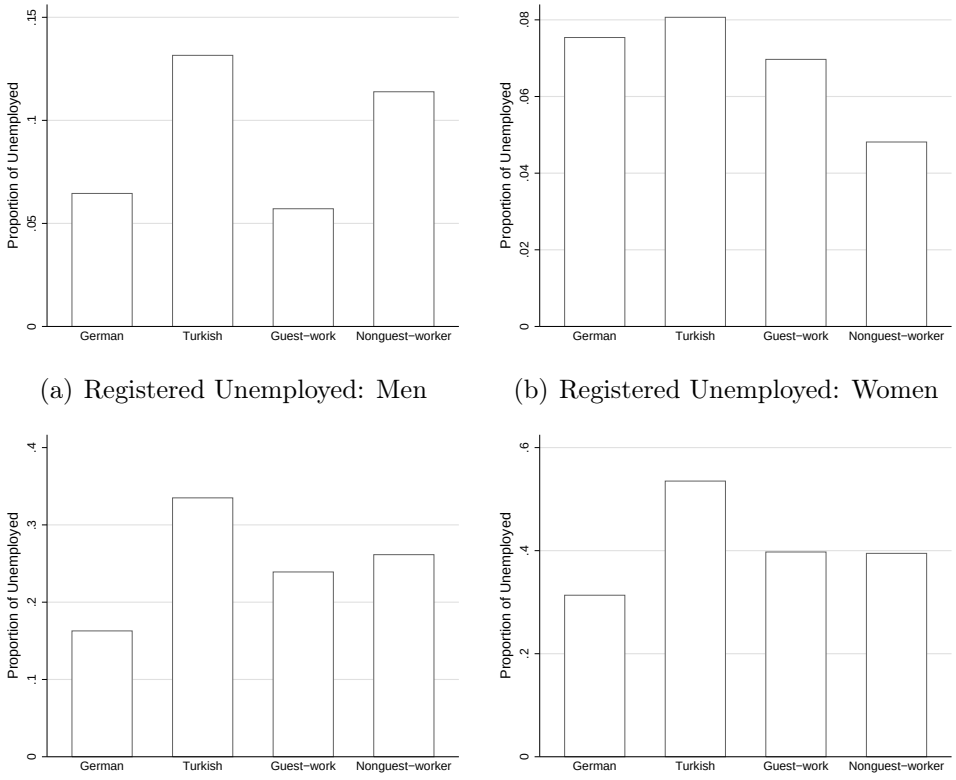


(a) Men (b) Women

Figure 1: Distribution of hourly wages measured in €

In Fig. 2 the proportion of individuals not working for two different proxies of unemployment is presented. Figs. 2(a) and 2(b) use the indicator *registered as unemployed at the Federal Agency of Employment*, while the measure used for constructing Figs. 2(c) and 2(d) is based on the survey question *currently working*. Both measures suggest that the largest proportion of unemployed individuals are Turkish second generation immigrants and that this difference is more pronounced for Turkish men than for Turkish women. With respect to the official employment indicator, guest-worker men and women that exclude the Turkish, have a smaller or equal proportion of individuals in unemployment than German natives. For non guest-worker men and women the picture is different. Whereas women from non guest-worker backgrounds have a substantially lower proportion of individuals in unemployment than German natives, men have a larger proportion. There is, however, a substantial proportion of women from other foreign backgrounds

that are currently not working (Fig. 2(d)). For this measure differences between German natives and second generation immigrants appear to be the largest. Tables 12 and 13 in Appendix A provide an overview of the summary statistics of all variables used in the analysis.



(a) Registered Unemployed: Men

(b) Registered Unemployed: Women

(c) Not working: Men

(d) Not working: Women

Figure 2: Proportion of men and women in unemployment

Whether a dynamic specification is appropriate is tested with an approach derived by Wooldridge (2002) for panel data and implemented by Drukker (2003) in STATA. For both men and women, I reject the Null Hypothesis of no first order autocorrelation at the 1 % significance level for wages and unemployment probabilities (Results are provided upon request)[10].

5.2 Estimation Results

In the next two sub-sections I present pooled OLS and maximum likelihood estimates of the parameters β_g, ψ_0, ψ_1, and ρ of Eqs. (11), (14) and Eq. (15). Full results of the

[10]The test is suitable for linear models only, however. Therefore, the unemployment equation is tested with a linear probability model.

baseline model are presented in the Appendix.

Hourly Wages

The first and second columns of Tables 4 and 5 contain estimates of a model that includes indicators of the second generation only. Model (1) assumes that the second generation is a homogeneous group and Model (2) assumes that hourly wages differ between Turkish, Yugoslav, Greek, Italian, Spanish, and non guest-worker immigrants.

In the raw data, men of the second generation earn 6.83 € or 28.4 % less than German natives, who earn 9.74 €[11]. The same holds approximately for Turks and Greek. Yugoslav men earn on average 37 % less than German natives, while Spanish men experience the smallest wage difference of - 25 %.

Female second generation immigrants earn on average 5.95 € per hour or 26 % less than German natives, who earn 8.06 € per hour. For all guest-worker women (except Turkish), the wage differences relative to German natives are similar to those of guest-worker men, except that they are slightly smaller in magnitude. For both male and female second generation immigrants from non guest-worker backgrounds there are no statistically significant differences in the raw data.

The subsequent models of Tables 4 and 5 report the results of adding regressors to the baseline model. Model (3) adds age-group and education dummy variables and Model (4) adds household characteristics, marital status, locational and time dummy variables (Eq. (11)). Model (5) re-estimates Model (4) by excluding all part-time employed and Model (6) re-estimates Model (4) with a linear random effects Mundlak specification (Eq. (13)). Model (7) re-estimates Model (4) by adding proxies of socioeconomic status of the parents such as the father's employment status when the individual was aged 15 and the religion of the parents. This model is compared then to Model (8) in which intergenerational transmission is controlled for with dynamics and a Mundlak specification of the error term.

For men, all significant earnings disadvantages disappear once controlling for age and education, they even turn positive for most second generation sub-groups (Model (3)).

[11] According to Halvorsen and Palmquist (1980), the interpretation of dummy variables in a semi-logarithmic linear regression is $\exp(\hat{\beta}) - 1$, e.g. we have for the second generation men in the raw data (Model (1)): $\Delta_{sg} = \exp(-.335) - 1 = -.284$.

Any hourly wage advantage or disadvantage can be explained by observable factors such as time and locational effects, household characteristics and marital status for almost all sub-groups. Only Spanish men earn robustly higher wages that are, depending on the model, 7 to 13 % greater than those of German natives. These higher wages are mainly earned in the part-time sector (Model (5)).

Controlling for unobserved heterogeneity with a Mundlak specification (Model (6)) yields similar results as dropping all part-time employed for the model (Model (5)), at least in terms of sign of the coefficients. Also, controlling for the parents's socioeconomic status in Model (7) yields results close to those of a dynamic Mundlak specification (Model (8)) for Spanish men.

Table 4: Estimated hourly wage for men

	POLS HOMOG RAW DATA	POLS HETEROG RAW DATA	POLS AGE EDU	POLS ALL VARS	POLS ONLY FTE	RE MUND ALL	POLS PARENT SES	RE MUND DYN
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Second generation	-.335***							
	(.024)							
Turkish		-.322***	.173***	.007	.008	.008	.026	.006
		(.037)	(.027)	(.026)	(.026)	(.030)	(.041)	(.025)
Yugoslav		-.456***	.132***	-.025	-.005	.010	-.016	.062*
		(.066)	(.040)	(.039)	(.043)	(.046)	(.040)	(.037)
Greek		-.360***	.094**	.007	.016	.061	.029	.012
		(.064)	(.045)	(.045)	(.051)	(.047)	(.045)	(.035)
Italian		-.384***	.128***	.022	.036	.058	.013	.042
		(.046)	(.034)	(.032)	(.033)	(.039)	(.032)	(.026)
Spanish		-.289***	.136***	.076**	.033	.159***	.066*	.068*
		(.067)	(.035)	(.034)	(.033)	(.050)	(.034)	(.036)
Non guest-worker		.018	.008	-.069	-.005	-.064	-.073	.005
		(.096)	(.082)	(.067)	(.057)	(.075)	(.066)	(.049)
Hourly wage in t-1								.348***
								(.005)
Hourly wage in t1								.098***
								(.006)
Const.	2.277***	2.277***	1.006***	1.515***	1.890***	1.505***	1.536***	.955***
	(.006)	(.006)	(.021)	(.026)	(.033)	(.105)	(.027)	(.079)
Obs.	59244	59244	59244	58316	48142	58316	58315	33743
R^2	.014	.016	.456	.545	.437		.546	
F statistic	201.467	38.075	1165.472	494.119	255.86		433.305	
ρ						.504		.187

Table 4 reports the coefficients of regressing the log of hourly wage of men on a set of second generation indicators. Second generation immigrants are foreign nationals who arrived in Germany no later than 6 years of age. Model (1) includes an indicator for the second generation as a whole and Model (2) distinguishes between Turkish, Yugoslav, Italian, Greek, and Spanish (including Portuguese) and non guest-worker. Both models exclude all control variables. Model (3) includes age-interval and education dummy variables, Model (4) includes all control variables of the prior models plus marital status, the number of children in the household, location, and time dummy variables. Model (5) re-estimates Model (4) by excluding all individuals who are partially employed, where *FTE* stands for *full time employment*. Model (6) re-estimates Model (5) with a linear random effects Mundlak specification. Model (7) re-estimates Model (4) by adding proxies for parents' socioeconomic status. Model (8) re-estimates Model (4) by adding lagged values of the dependent variable and its initial condition. Robust standard errors are reported in parentheses. ρ measures the fraction of variance due to r_i. * 10 %, ** 5 %, *** 1 % significance level.

Yugoslav men attract higher wages by more than 6 % than German natives in the dynamic model, in which both time-invariant unobserved heterogeneity and persistence are controlled for. Whether the dynamics in Model (8) are justified is investigated with a Wald Test, testing for the Null Hypothesis of statistical insignificance of the sum of the two coefficients on the lagged value of the hourly wage and its initial condition. I reject the Null Hypothesis at the 1 % level of statistical significance (Results are provided upon request)[12]. The large differences in ρ between Models (6) and (8) suggest that a large fraction of the total variation in the composite error term is explained by the dynamic nature of wages ($>$ 30 % points).

I also test for the presence of random effects in the model with a Breusch-Pagan Lagrange Multiplier (BPLM) test. For both Models (6) and (8), the Null Hypothesis of a zero variance of the individual fixed effect ($H_0 : var(r_i) = 0$) can be rejected at the 1 % significance level (Results are provided upon request).

The results on the estimated hourly wage of women, reported in Table 5, tell almost an identical story. Once controlling for age, education, household characteristics, locational effects, and time effects, none of the groups of second generation women faces a statistical significant disadvantage. Even though the coefficients are statistically not significant (due to large standard errors), the wage coefficients of Yugoslav, Italian, and Spanish women are robustly positive across all model specifications. Particularly in the dynamic specification, the estimated hourly wages for these three groups are prominently greater (For Italian women they are 10 % greater and statistically significant at the 5% level).

Similarly as for men, the Null Hypothesis of statistical insignificance of the sum of the coefficients on the lagged value of the hourly wage and its initial condition is rejected at the 1 % level (Results are provided upon request)[13].

[12]The Null Hypothesis states: $H_0 : \psi_0 + \psi_1 = 0$. The dynamic model reveals that current wages depend strongly on the past period's wage (by approximately 1/3) and on the initial wage (by approximately 1/10). In sum, these two coefficients add up to 0.43, which is slightly less, but similar in magnitude, than the common result of 0.5 found in the literature on the wage curve (Blanchflower and Oswald, 2005).

[13]For women, dependence on initial values of the hourly wage plays a slightly greater role (14 %) in determining current wages than for men (10 %), but past period's wages have a slightly smaller influence on current wages (less than 30 %).

Table 5: Estimated hourly wage for women

	POLS HOMOG RAW DATA	POLS HETEROG RAW DATA	POLS AGE EDU	POLS ALL VARS	POLS ONLY FTE	RE MUND ALL	POLS PARENT SES	RE MUND DYN
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Second generation	-.304***							
	(.025)							
Turkish		-.313***	.077*	-.014	.023	.063	-.035	-.009
		(.035)	(.040)	(.039)	(.033)	(.039)	(.066)	(.046)
Yugoslav		-.334***	.122***	.026	.003	.049	.031	.070
		(.047)	(.036)	(.035)	(.037)	(.052)	(.036)	(.053)
Greek		-.292***	.042	-.023	-.012	.047	-.007	-.061
		(.074)	(.051)	(.048)	(.046)	(.054)	(.049)	(.055)
Italian		-.303***	.068*	.014	.034	.066	.021	.091**
		(.054)	(.041)	(.041)	(.052)	(.045)	(.041)	(.041)
Spanish		-.362***	.029	.018	.033	.004	.032	.065
		(.123)	(.090)	(.087)	(.085)	(.075)	(.090)	(.071)
Non guest-worker		-.078	.034	-.118	-.105	-.097	-.116	-.046
		(.132)	(.075)	(.073)	(.075)	(.096)	(.073)	(.100)
Hourly wage in t-1								.285***
								(.006)
Hourly wage in t1								.143***
								(.007)
Const.	2.088***	2.088***	.938***	1.380***	1.693***	1.706***	1.382***	1.038***
	(.006)	(.006)	(.020)	(.028)	(.036)	(.107)	(.030)	(.103)
Obs.	50,535	50,535	50,535	49,469	25,808	49,469	49,469	24,249
R^2	.01	.011	.316	.392	.372		.393	
F statistic	147.507	29.89	672.368	293.955	128.788		254.685	
ρ						.484		.217

Table 5 reports the coefficients of regressing the log of hourly wage of women on a set of second generation indicators. Second generation immigrants are foreign nationals who arrived in Germany no later than 6 years of age. Model (1) includes an indicator for the second generation as a whole and Model (2) distinguishes between Turkish, Yugoslav, Italian, Greek, and Spanish (including Portuguese) and non guest-worker. Both models exclude all control variables. Model (3) includes age-interval and education dummy variables, Model (4) includes all control variables of the prior models plus marital status, the number of children in the household, location, and time dummy variables. Model (5) re-estimates Model (4) by excluding all individuals who are partially employed, where *FTE* stands for *full time employment*. Model (6) re-estimates Model (5) with a linear random effects Mundlak specification. Model (7) re-estimates Model (4) by adding proxies for parents' socioeconomic status. Model (8) re-estimates Model (4) by adding lagged values of the dependent variable and its initial condition. Robust standard errors are reported in parentheses. ρ measures the fraction of variance due to r_i. * 10 %, ** 5 %, *** 1 % significance level.

Decomposition of Wage Differentials

In Figures 3 and 4 I illustrate the differences in predicted hourly wages between German natives and each second generation sub-group separately (*Differences*) and the proportions to which these differences can be attributed to differences in observed characteristics (*Endowments*), differences in coefficients (*Coefficients*), or differences of interactions between coefficients and characteristics (*Interactions*). Full results are presented in the Appendix.

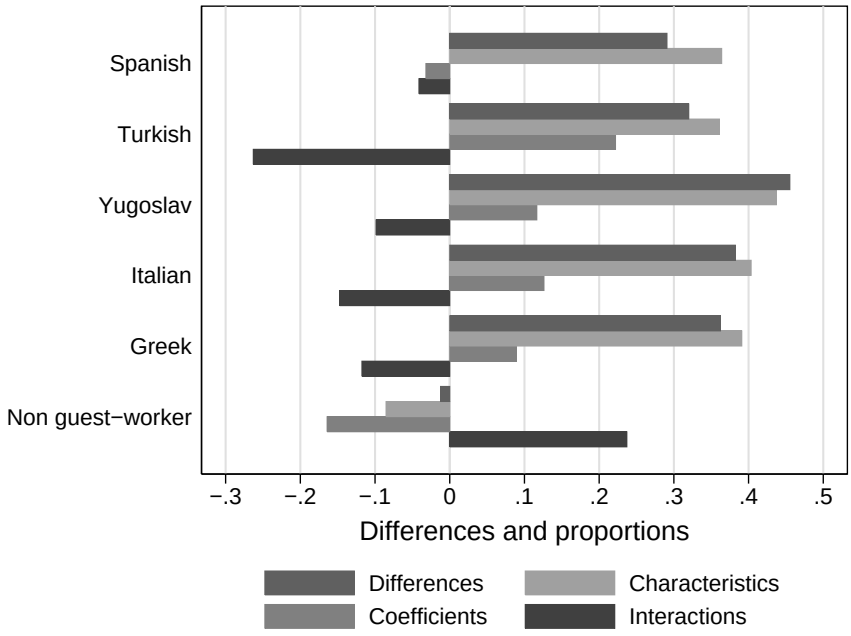


Figure 3: Threefold Oaxaca-Blinder decomposition of hourly wages for men

For both men and women, the decomposition analysis tells a straight-forward story. For all sub-groups the estimated differences (first horizontal bar for each sub-group) are positive, that means the predicted wages of German natives are larger than the predicted wages of all second generation immigrants (except for non guest-worker men), when controlling for the full set of observable characteristics, including the parents' socioeconomic status. Most differences in predicted wages between German natives and any sub-group of second generation immigrants are attributed to observable characteristics (second horizontal bar). Only for non guest-worker men and women differences in the coefficients (third horizontal bar) explains a larger part of the estimated difference in wages. For both Turkish men and women, the interaction of coefficients and characteristics, i.e. the

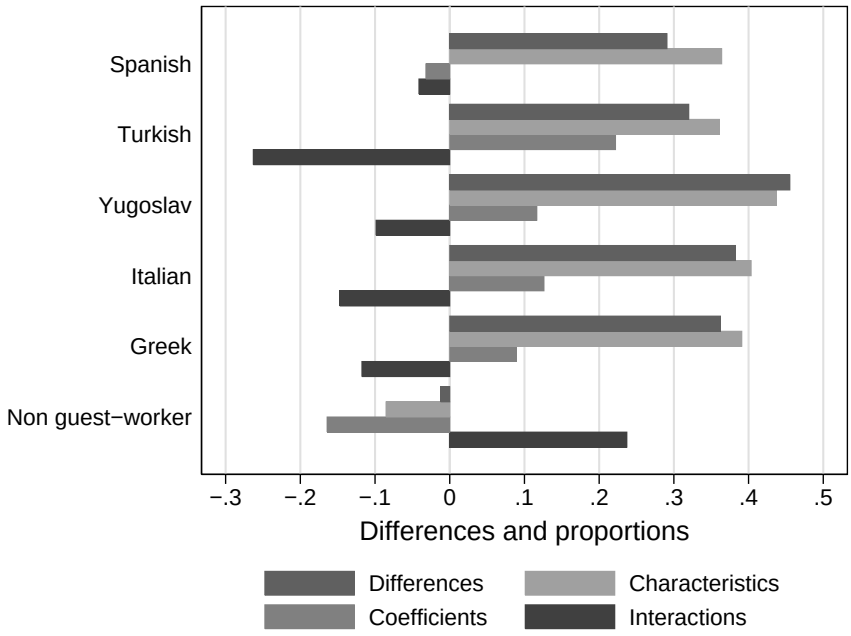


Figure 4: Threefold decomposition of hourly wages for women

simultaneous effect of observable and unobservable characteristics plays a large role in determining wage differentials.

Overall, these results suggest that a large fraction of the difference in wages between sub-groups of second generation immigrants and German natives are driven by differences in observable characteristics such as age, education, and socioeconomic characteristics of the parents. Unobservable factors, such as cognitive ability, motivation, degree of migration background and neighbourhood effects are less relevant in explaining the wage differences.

Unemployment

The estimates of the raw differentials in unemployment risks between second generation immigrants and German native men and women are reported in the first two columns of Tables 6 and 7. The second generation, considered as a homogeneous group, is 2.4 % more likely to be registered unemployed than German natives. However, this small figure is mainly the result of a relatively large probability to be registered unemployed for Turkish men (almost 7 %) and a relatively small probability to be unemployed by Spanish men (- 2.5 %). For all other nationalities, differences are not statistically significant at any

conventional significance level. For women the estimated results suggest no statistically significant differences in the risk of unemployment between second generation immigrants and German natives in the raw data.

The subsequent columns of the same tables report the results of unemployment probabilities when comparing the comparable, i.e. adding control variables to the estimated model. For men, it is still the Turkish second generation who faces a higher risk of unemployment of 3 to 7 % across all model specifications. A large fraction of 25 % to 50 % of this higher risk can be explained by time-invariant unobservable heterogeneity and dynamics in the data (Model (6) and (8)). Interestingly, the dynamic random effects model specification, which does not control for parents' socioeconomic characteristics, yields the same unemployment probability of approximately 3.5 % for Turkish men as the pooled OLS model that controls directly for these factors.

A similar trend emerges for second generation immigrants from non guest-worker backgrounds, who face a higher risk of unemployment between 3 and 5 % that is only statistically significant in Model (8). Yugoslav men are only marginally more likely to be registered unemployed of approximately 2 %, and these differences are only prominent once unobserved time-invariant factors and dynamics are controlled for.

The smaller risk of Spanish and Italian second generation men vis-à-vis German natives of 2 to 3 % and 1 to 2 %, respectively cannot be explained by age or human capital endowment, but by other observable characteristics such as household variables, time trends, and locational dummy variables (Model 5)). Greek men face similar unemployment risks as German natives.

Table 6: Unemployment risks for men

	PP HOMOG RAW DATA	PP HETEROG RAW DATA	PP Age EDU	PP AGE EDU	PP ALL VARS	REP MUND-LAK	PP PARENTS SES	REP DYNAMICS MUNDLAK
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Second generation	0.024							
	(0.007)***							
Turkish		0.067	0.070	0.029	0.060	0.045	0.036	0.034
		(0.013)***	(0.014)***	(0.010)***	(0.012)***	(0.007)***	(0.015)**	(0.006)***
Yugoslav		0.010	0.014	-0.005	0.017	0.019	0.007	0.017
		(0.015)	(0.016)	(0.012)	(0.014)	(0.011)*	(0.013)	(0.009)*
Greek		-0.001	0.001	-0.006	0.014	0.014	0.001	0.011
		(0.014)	(0.015)	(0.012)	(0.011)	(0.012)	(0.010)	(0.010)
Italian		-0.013	-0.011	-0.025	-0.003	0.003	-0.002	0.001
		(0.010)	(0.010)	(0.007)***	(0.008)	(0.010)	(0.008)	(0.009)
Spanish		-0.025	-0.023	-0.030	-0.011	0.003	-0.012	-0.001
		(0.009)***	(0.009)**	(0.007)***	(0.009)	(0.013)	(0.008)	(0.012)
Non guest-worker		0.050	0.052	0.043	0.056	0.028	0.053	0.028
		(0.047)	(0.048)	(0.039)	(0.036)	(0.019)	(0.033)	(0.016)*
Unemployment in t-1								0.061
								(0.002)***
Unemployment in t1								0.060
								(0.004)***
Observations	76,778	76,778	76,778	76,778	75,588	75,588	75,585	65,008
ρ						.498		.298

Table 6 reports the marginal effects of augmentations of a model that regresses the unemployment probability on a set of second generation indicators for men. Second generation immigrants are foreign nationals who arrived in Germany no later than 6 years of age. Model (1) includes the second generation as one homogeneous group and Model (2) distinguishes between Turkish, Yugoslavs, Italian, Greek, Spanish, and non guest-worker children. Both models add no control variables. Model (3) adds age-interval indicators, Model (4) adds education indicators, Model (5) adds household specific variables such as marital status and persons in household, location, and time dummy variables. Model (6) re-estimates Model (5) with a random effects probit model with a Mundlak specification of the error term. Model (7) re-estimates Model (5) adding parental socioeconomic background variables, and Model (8) adds a lagged value and its initial value of the dependent variable. ρ measures the fraction of variance due to r_i. Marginal effects for Model (4) and (5) are calculated from coefficients adjusted by $\sqrt{1-\hat{\rho}}$ and using the Delta method in STATA. * 10 %, ** 5 %, *** 1 % significance level.

For women, once comparing the comparable in Table 7, it is only Yugoslav second generation immigrants who are estimated to face a greater risk of being unemployed of 2 to 6 %, depending on the model assumptions. A great proportion of the higher unemployment risk reported in Model (5) is attributed to unobserved heterogeneity, as the marginal effects in Model (6) and (8) are reduced by 1/3 and 1/2, respectively. These results are surprising, given that Yugoslav women were estimated to attract higher wages than comparable German natives. The share of Yugoslav women who find employment in the first place is significantly lower than the proportion of German women.

Turkish women also face a higher unemployment risk as well, but of a smaller magnitude than Yugoslav women (1.5 to 2%). An interesting point to note is that the statistically significant small differences in unemployment for the Turkish women is evident only after controlling for the full set of observable characteristics, unobservable factors or dynamics. However, they can be explained by the socioeconomic characteristics of the parents (Model (7)).

The extent to which unobserved heterogeneity plays a role in Models (6) and (8) can be interpreted with the help of ρ. In Model (6) the fraction of the variance determined by time-invariant unobserved heterogeneity accounts for approximately 50 % for both men and women. Once controlling for the dynamics of wages, this proportion is reduced to 29 % for men and 26 % for women suggesting that controlling for dynamics in the model picks up some additional form of unobservable heterogeneity. The Null Hypothesis of a joint insignificance of the sum of the persistence parameter and the initial condition is rejected at a 1 % significance level for both men and women (Results are provided upon request)[14].

[14]The persistence and the initial condition parameters in Model (5) are highly statistically significant for both men and women and averages in sum to .075 and 0.11 respectively.

Table 7: Unemployment risks for women

	PP HOMOG RAW DATA	PP HETEROG RAW DATA	PP Age EDU	PP AGE EDU	PP ALL VARS	REP MUND-LAK	PP PARENTS SES	REP DYNAMICS MUNDLAK
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Second generation	-0.003							
	(0.006)							
Turkish		0.004	0.012	-0.007	0.019	0.019	0.011	0.015
		(0.010)	(0.010)	(0.009)	(0.010)*	(0.009)**	(0.012)	(0.008)*
Yugoslav		0.018	0.026	0.021	0.064	0.045	0.057	0.030
		(0.017)	(0.018)	(0.017)	(0.020)***	(0.013)***	(0.020)***	(0.010)***
Greek		-0.012	-0.006	-0.015	0.018	0.009	0.009	0.011
		(0.015)	(0.016)	(0.013)	(0.018)	(0.015)	(0.016)	(0.012)
Italian		-0.020	-0.017	-0.029	0.002	-0.002	0.001	-0.002
		(0.012)	(0.013)	(0.010)***	(0.013)	(0.013)	(0.013)	(0.011)
Spanish		-0.005	0.001	-0.008	0.018	0.029	0.017	0.026
		(0.023)	(0.024)	(0.020)	(0.023)	(0.019)	(0.024)	(0.016)
Non guest-worker		-0.005	-0.025	-0.024	0.009	0.016	0.008	0.027
		(0.008)	(0.023)	(0.022)	(0.028)	(0.026)	(0.027)	(0.021)
Unemployment in t-1								0.083
								(0.002)***
Unemployment in t1								0.069
								(0.004)***
Observations	78,444	78,444	78,447	78,447	76,836	76,836	76,836	65,998
ρ						.503		.261

Table 7 reports the marginal effects of augmentations of a model that regresses the unemployment probability on a set of second generation indicators for women. Second generation immigrants are foreign nationals who arrived in Germany no later than 6 years of age. Model (1) includes the second generation as one homogeneous group and Model (2) distinguishes between Turkish, Yugoslavs, Italian, Greek, Spanish, and non guest-worker children. Both models add no control variables. Model (3) adds age-interval indicators, Model (4) adds education indicators, Model (5) adds household specific variables such as marital status and persons in household, location, and time dummy variables. Model (6) re-estimates Model (5) with a random effects probit model with a Mundlak specification of the error term. Model (7) re-estimates Model (5) adding parental socio-economics background variables, and Model (8) adds a lagged value and its initial value of the dependent variable. ρ measures the fraction of variance due to r_i. Marginal effects for Model (4) and (5) are calculated from coefficients adjusted by $\sqrt{1-\hat{\rho}}$ and using the Delta method in STATA. * 10 %, ** 5 %, *** 1 % significance level.

Decomposition of Unemployment Risk Differentials

In Figures 5 and 6 the differences in estimated unemployment risks between German natives and sub-groups of second generation immigrants are illustrated for a model that controls for all observable characteristics including parents' socioeconomic background. Decomposition results could not be obtained for Spanish and non guest-worker women, most likely due to the small sample size of these two groups.

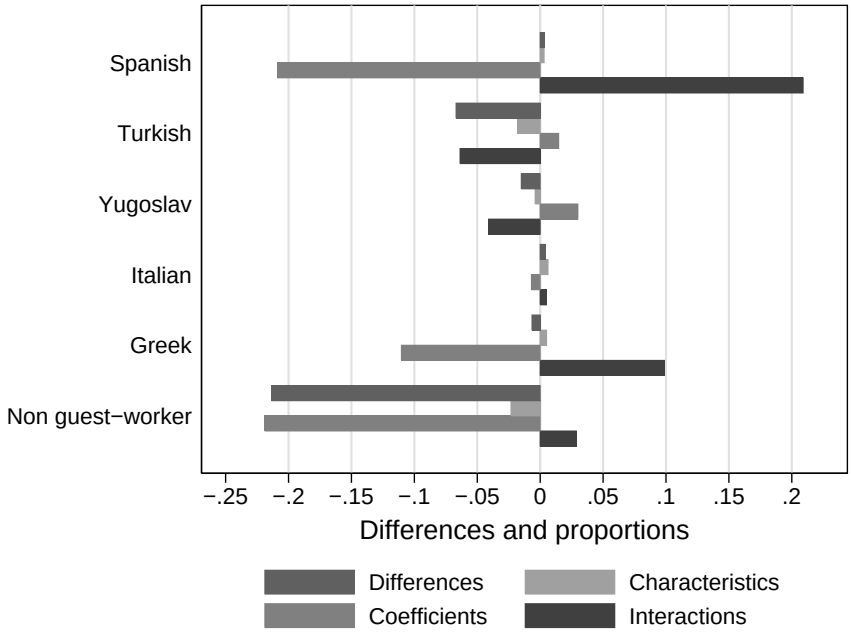


Figure 5: Threefold Oaxaca-Blinder decomposition of unemployment risks for men

The first horizontal bar of each sub-group indicates the predicted differences in unemployment probabilities between German natives and a sub-group of second generation immigrants. For almost all sub-groups, men and women alike, these differences are negative, indicating that German natives are predicted to be less likely unemployed than second generation immigrants. Only Spanish and Italian men are less likely to be unemployed than German natives.

The second, third, and fourth bar of each sub-group report the proportions to which the differences can be attributed to differences in observed characteristics (*Endowments*), differences in coefficients (*Coefficients*), and differences of interactions between coefficients and characteristics (*Interactions*).

In contrast to the decomposition of hourly wages, all unemployment differentials are

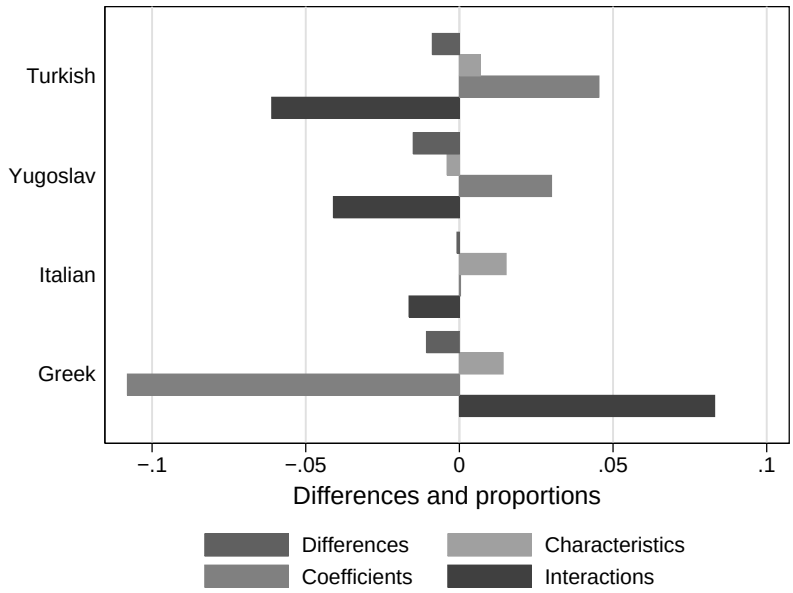


Figure 6: Threefold decomposition of unemployment risks

driven by differences in coefficients and differences in interaction effects for both men and women. This suggests that it is mainly unobserved differences, i.e. differences in ability, motivation, neighbourhood effects, and the degree of migration background that drives unemployment risk differentials. Given these differences in underlying causes, my results propose that wage earnings differences are driven by different factors than the probability of becoming unemployed.

6 Robustness Checks

In this section I test whether the differences in wage earning capacity for Spanish men and Yugoslav women (who both earn more than their German native counterparts) and in unemployment risks for Turkish and non guest-worker men and Yugoslav and Turkish women (who have higher risks than German natives) are robust. Specifically, the estimated differences of the preferred model that controls for all observable characteristics including socioeconomic background of the parents (Model 1) are tested against the use of probability weights[15] (Model 2), the definition of the outcome variable (Model 3), the definition of the second generation (Model 4), and a sample that excludes all second

[15]Probability weights are usually used in the GSOEP to account for the over-sampling of foreigners in the data set. E.g. Fertig and Schurer (2007) and Brenner (2007) use probability weights and discuss their limitations in econometric modelling.

generation immigrants that acquired the German citizenship[16] (Model 5).

As probability weights can only be applied to pooled models, these are used as basis for the robustness checks. Since the pooled model that includes all socioeconomic characteristics of the parents yields similar results as a dynamic random effects Mundlak model, I choose the former model as benchmark case.

Table 8: Robustness checks on estimated hourly wage for men

	Pooled OLS including all variables plus parents' socioeconomic status				
	FULL MODEL	PWEIGHTS	ALT DV	STRICT	CHANGE
	(1)	(2)	(3)	(4)	(5)
Turkish	.026	-.030	.018	-.037	.009
	(.041)	(.047)	(.043)	(.050)	(.044)
Yugoslav	-.016	-.054	-.016	-.013	-.075
	(.040)	(.045)	(.037)	(.052)	(.053)
Greek	.029	-.021	.042	-.008	.026
	(.045)	(.066)	(.050)	(.054)	(.051)
Italian	.013	.0004	.015	.010	.025
	(.032)	(.039)	(.033)	(.041)	(.032)
Spanish	.066*	.030	.078**	.051	.066*
	(.034)	(.056)	(.032)	(.048)	(.034)
Non guest-worker	-.073	-.102	-.073	-.142*	.034
	(.066)	(.084)	(.063)	(.077)	(.095)
Const.	1.536***	1.532***	1.717***	1.521***	1.529***
	(.027)	(.039)	(.028)	(.028)	(.027)
Obs.	58315	56083	63018	56756	57533
R^2	.546	.515	.615	.546	.546
F statistic	433.305	217.674	755.79	425.804	427.417

Table 8 reports the robustness checks on the results of the estimated hourly wage of men when using the preferred pooled OLS model. This model controls for age, human capital, father's socioeconomic status and religion, family characteristics, time and locational dummies (results omitted). Model (1) corresponds to the preferred model (FULL MODEL), Model (2) applies probability weights to account for the over-sampling of foreigners in the data set (PWEIGHTS), Model (3) uses an alternative measure of the hourly wage (constructed from gross annual earnings) (ALT DV), Model (4) excludes all second generation immigrants who were not born in Germany (STRICT), and Model (5) excludes all second generation immigrants who acquired the German citizenship (90 individuals). Robust standard errors are reported in parentheses. * 10 %, ** 5 %, *** 1 % significance level.

Table 8 reports the changes in the coefficients of differences in hourly wages for men. Overall, the picture of hourly wage differences between second generation men and comparable German natives does not change. The Spanish earn a higher hourly wage than German natives, a difference which is even more pronounced when using the an alternative income measure (Almost 8 %)[17]. Also, the trend of lower hourly wages for non guest-workers is most pronounced for those who were born in Germany (- 14 %). For women, the robustness checks on estimated hourly wages do not alter the main conclusions either. Yugoslav women are estimated to earn higher hourly wages than comparable

[16]The sample includes 90 men and 103 women who obtained German citizenship in due course of the panel.

[17]This measures individual annual labour earnings constructed from the salary obtained through main job.

Table 9: Robustness checks on estimated hourly wage for women

	Pooled OLS including all variables plus parents' socioeconomic status				
	FULL MODEL	PWEIGHTS	ALT DV	STRICT	CHANGE
	(1)	(2)	(3)	(4)	(5)
Turkish	-.027	-.080	-.033	-.059	.061
	(.065)	(.068)	(.066)	(.080)	(.048)
Yugoslav	.069**	.079	.037	.072**	.089**
	(.033)	(.049)	(.036)	(.036)	(.045)
Greek	.076	.010	-.006	.019	.081
	(.051)	(.086)	(.049)	(.054)	(.056)
Italian	.034	.036	.027	.041	.045
	(.040)	(.045)	(.041)	(.041)	(.043)
Spanish	.071	.027	.033	.026	.069
	(.094)	(.074)	(.091)	(.100)	(.102)
Non guest-worker	-.056	-.124**	-.100	.004	-.019
	(.064)	(.058)	(.072)	(.084)	(.063)
Const.	1.501***	1.455***	1.381***	1.487***	1.487***
	(.030)	(.046)	(.030)	(.031)	(.031)
Obs.	53,076	50,642	49,467	52,229	52,308
R^2	.504	.464	.393	.504	.503
F statistic	498.101	231.481	254.793	492.764	492.447

Table 9 reports the robustness checks on the results of the estimated hourly wage of women when using the preferred pooled OLS model. This model controls for age, human capital, father's socioeconomic status and religion, family characteristics, time and locational dummies (results omitted). Model (1) corresponds to the preferred model (FULL MODEL), Model (2) applies probability weights to account for the over-sampling of foreigners in the data set (PWEIGHTS), Model (3) uses an alternative measure of the hourly wage (constructed from gross annual earnings) (ALT DV), Model (4) excludes all second generation immigrants who were not born in Germany (STRICT), and Model (5) excludes all second generation immigrants who acquired the German citizenship (103 individuals). Robust standard errors are reported in parentheses. * 10 %, ** 5 %, *** 1 % significance level.

German natives, and this advantage is slightly more pronounced when excluding all Yugoslav women who acquired German citizenship. Greek, Italian, and Spanish women are estimated robustly to earn higher wages across all model specification (even though the differences are not statistically significant). Estimated wages for Turkish women have consistently a negative sign across all robustness tests. The negative sign, however, seems to be driven by those Turkish second generation immigrants who changed nationality.

Table 10 and 11 report the robustness checks for the estimated differentials in unemployment risks for men and women. Turkish men face a statistically significant greater risk of being registered unemployed (3% to 4 %), independent of whether the sample includes only individuals strictly born in Germany or excludes individuals who acquired German citizenship. Turkish men are more likely to be registered unemployed, but they are not more likely to be not working.

Table 10: Robustness checks on estimated risk of unemployment for men

	Pooled probit including all variables plus parents' socioeconomic status				
	FULL MODEL	PWEIGHTS	ALT DV	STRICT	CHANGE
	(1)	(2)	(3)	(4)	(5)
Turkish	0.036	0.031	0.003	0.042	0.032
	(0.015)**	(0.023)	(0.023)	(0.019)**	(0.015)**
Yugoslav	0.007	0.012	-0.004	0.012	0.036
	(0.013)	(0.019)	(0.020)	(0.016)	(0.022)
Greek	0.001	-0.012	0.018	0.002	-0.001
	(0.010)	(0.011)	(0.027)	(0.011)	(0.010)
Italian	-0.002	-0.003	-0.061	-0.002	-0.004
	(0.008)	(0.014)	(0.012)***	(0.011)	(0.008)
Spanish	-0.012	-0.029	-0.049	0.002	-0.013
	(0.008)	(0.007)***	(0.016)***	(0.013)	(0.008)
Non guest-worker	0.053	0.112	0.065	0.080	0.040
	(0.033)	(0.045)**	(0.051)	(0.048)*	(0.038)
Observations	75585	72606	76057	73591	74528

Table 10 reports the robustness checks to the estimated unemployment risks obtained from a model that controls for all variables plus socioeconomic characteristics of the parents. Model (1) corresponds to the preferred model (FULL MODEL), Model (2) applies probability weights to account for the over-sampling of foreigners in the data set (PWEIGHTS), Model (3) uses an alternative measure of unemployment, i.e. *currently not working* (ALT DV), Model (4) excludes all second generation immigrants who were not born in Germany (STRICT), and Model (5) excludes all second generation immigrants who acquired the German citizenship (103 individuals). Robust standard errors are reported in parentheses. * 10 %, ** 5 %, *** 1 % significance level.

Table 11: Robustness checks on estimated risk of unemployment for women

	Pooled probit including all variables plus parents' socioeconomic status				
	FULL MODEL	PWEIGHTS	ALT DV	STRICT	CHANGE
	(1)	(2)	(3)	(4)	(5)
Turkish	0.011	0.017	0.001	0.025	0.005
	(0.012)	(0.018)	(0.035)	(0.016)	(0.013)
Yugoslav	0.057	0.052	-0.020	0.046	0.058
	(0.020)***	(0.036)	(0.032)	(0.022)**	(0.033)*
Greek	0.009	0.001	0.006	0.001	0.012
	(0.016)	(0.025)	(0.042)	(0.018)	(0.017)
Italian	0.001	-0.011	-0.055	-0.002	0.007
	(0.013)	(0.014)	(0.028)**	(0.014)	(0.014)
Spanish	0.017	0.104	-0.002	0.004	0.011
	(0.024)	(0.065)	(0.060)	(0.027)	(0.025)
Non guest-worker	0.008	-0.048	0.060	0.008	0.006
	(0.027)	(0.006)***	(0.066)	(0.027)	(0.029)
Observations	76,836	73,776	77,340	75,369	75,608

Table 11 reports the robustness checks to the estimated unemployment risks obtained from a model that controls for all variables plus socioeconomic characteristics of the parents. Model (1) corresponds to the preferred model (FULL MODEL), Model (2) applies probability weights to account for the over-sampling of foreigners in the data set (PWEIGHTS), Model (3) uses an alternative measure of unemployment, i.e. *currently not working* (ALT DV), Model (4) excludes all second generation immigrants who were not born in Germany (STRICT), and Model (5) excludes all second generation immigrants who acquired the German citizenship (103 individuals). Robust standard errors are reported in parentheses. * 10 %, ** 5 %, *** 1 % significance level.

Similarly, non guest-workers are estimated to face a greater risk of being registered as unemployed by 8 % to 11 % when excluding all individuals who were not born in Germany or when applying probability weights in the estimation. Yugoslav and Greek men are estimated to have no greater risk than German natives to be registered unemployed,

independent of the definition of the sample or the outcome variable. Interestingly, Italian and Spanish men, even though they do not differ in their probability to be registered unemployed, are less likely to stop working than comparable German natives (6 and 5 %, respectively).

For women, the robustness check does not alter conclusions: it is mainly Yugoslav women who have a greater risk of formal unemployment of 5 % to 6 %, but they are nt more likely to stop working. Italian women are less likely to stop working than comparable German natives, and that result is similar to those of Italian men.

7 Summary and Conclusion

This paper uses GSOEP data from 1984 to 2005 to examine the hypothesis that differences in productivity between German natives and second generation immigrants can be mainly explained by country of origin. Three principle findings of the analysis with the sample used suggest that (i) the second generation cannot be considered as one homogeneous group, as some nationalities perform better, equal or worse than German natives once departing from the raw data analysis, (ii) it is mainly Turkish and non guest-worker men and Yugoslav and, to a lesser extent, Turkish women who are most vulnerable in being formally unemployed, and therefore, being dependent on state benefits, and (iii) unobserved heterogeneity, which may represent an array of factors such as ability, motivation, neighbourhood effects, mixed marriage backgrounds, and socioeconomic characteristics of parents, plays a crucial role in explaining differences in unemployment risks but not in explaining wage differentials.

The relatively weak economic integration of Turkish children of immigrants mirror predictions made by Gestring et al. (2004), OECD (2005), Uhlendorff and Zimmermann (2006), and Worbs (2003). Differences in unemployment for this group can only be partially explained by observable characteristics such as age, education and socioeconomic status of the parents. A similar conclusion holds for non guest-worker men. In contrast, Turkish women do not seem to face the same high risk of unemployment as their gender counterpart.

More surprising is that Spanish second generation men, and to a lesser extent Spanish women, stick out most positively in the analysis of wages and unemployment risks. They

earn higher wages, are at a slightly lower risk of being unemployed, and are strictly less likely to stop working than comparable German natives. For women it is Yugoslav second generation immigrants who, consistently across all models, earn significantly higher wages than German natives. Greeks and Italians do not seem to differ largely from their German native counterparts, suggesting that Greeks and Italians of similar education levels, age and socioeconomic characteristics of the parents achieve similar results as Germans.

One may wonder why it is particularly the Spanish whose economic productivity is valued highly in the German labour market and why the Turkish struggle hard to integrate economically. The Spanish success story may be related to the strong political organisation of this ethnic minority in Germany. Thränhardt (1989) reports that the Spanish community adopted a pragmatic and effectively organised approach to the problems of Spanish guest-workers. These communities were instrumental in providing Spanish children with effective education institutions. Spanish immigrants asked early onwards for bilingual education, opted for full integration of their children into German schools and against special Spanish schools, and were pro-actively seeking for homework assistance programs. The Greek community was similarly well organised as the Spanish, even though the Greek community insisted more on maintaining Greek schools in Germany.

In contrast, the Turkish political organisation was geared towards a fundamentalist and radical orientatation. One dominant example is the Islamisches Kulturzentrum (Islamic Cultural Centre), which has more than 210 cultural centres throughout Germany and which is part of the fundamentalist movement of the Suleymanli sect. The cultural centre has a strong influence on the children of Turkish immigrants via its Koran courses organised throughout the country (Thränhardt, 1989, p. 16-17).

One may also wonder why the group of second generation immigrants of non guest-worker background performs relatively weakly in the local labour markets. One reason may be that the classification of this group comprises a variety of countries of origin, i.e. Western European, Eastern European, and Central Asian countries. The first group comprises countries such as Austria, France, Denmark, Great Britain, and USA, which are countries of similar religion, economic systems, and education institutions. The other group includes countries as diverse as Hungary, Czech Republic, Croatia, Bosnia and Herzegovina, Vietnam, and Azerbaijan, which differ in their religion, language origin, and economic systems. Hence, an additional degree of discrete heterogeneity may be implicit

to the group of non guest-workers. However, the small sample size does not allow for a further differentiation.

Another interesting outcome of my analysis is that most differences in hourly wages can be explained by observable characteristics such as age, education, marital status, time and locational effects, and parents' socioeconomic background. The same does not hold true for the differences in unemployment probabilities. The majority of the differences in unemployment risks is explained by either unobservable characteristics or by an interaction effect between observable and unobservable characteristics. This suggests that the determinants of wages are different from the determinants of unemployment. For the latter, one could think of personality characteristics such as tenacity, diligence, and perseverance that are instrumental in finding or keeping a job.

A Appendix

Table 12: Descriptive statistics for men

	All	Germans	SG	Turkish	GW	NGW
Hourly wage	10.5421	10.6842	8.435	8.4593	8.2354	11.1446
	(10.1008)	(10.1031)	(9.829)	(11.5703)	(8.8475)	(8.1352)
Gross labor income	2355.457	2396.717	1720.714	1656.1	1719.644	2317.137
	(1942.038)	(1973.963)	(1191.471)	(995.2444)	(1242.174)	(1651.947)
Unemployed (reg.)	.0662	.0645	.0884	.1315	.0571	.1139
	(.2486)	(.2457)	(.2839)	(.3381)	(.2321)	(.3182)
Not working	.1708	.1628	.2772	.335	.2391	.2615
	(.3764)	(.3692)	(.4476)	(.4721)	(.4266)	(.4402)
Months unemployed	.6307	.6139	.8476	1.237	.5497	1.1095
	(2.2327)	(2.199)	(2.6199)	(3.1873)	(2.0435)	(3.0009)
Time with firm	5.9386	6.2146	2.2721	1.8623	2.4057	4.2919
	(8.0416)	(8.1795)	(4.5285)	(4.3475)	(4.2877)	(7.3797)
Work hours (week)	43.7376	43.9389	40.6199	39.9644	41.1559	40.0556
	(10.0147)	(10.0537)	(8.823)	(8.5719)	(8.8678)	(9.516)
Age	33.2433	33.8776	24.8171	24.3226	24.5627	32.2792
	(9.6757)	(9.5879)	(6.306)	(5.586)	(6.0365)	(9.294)
Age-group 16 to 20	.1039	.0894	.2972	.3063	.3035	.1131
	(.3051)	(.2853)	(.457)	(.4611)	(.4598)	(.3172)
Age-group 21 to 30	.3157	.3004	.519	.5344	.5245	.3322
	(.4648)	(.4584)	(.4997)	(.4989)	(.4995)	(.4718)
Age-group 31 to 40	.3308	.3434	.164	.1544	.159	.3498
	(.4705)	(.4748)	(.3703)	(.3614)	(.3657)	(.4778)
Age-group 31 to 40	.3308	.3434	.164	.1544	.159	.3498
	(.4705)	(.4748)	(.3703)	(.3614)	(.3657)	(.4778)
Age-group 41 to 50	.2067	.221	.0174	.0049	.0114	.1767
	(.4049)	(.4149)	(.1307)	(.0702)	(.1062)	(.3821)
Age-group 51 to 60	.0429	.0459	.0024	0	.0017	.0283
	(.2026)	(.2093)	(.049)	(0)	(.0409)	(.166)
Max. 10 yrs school, no training	.0935	.0766	.3177	.4013	.2763	.1519
	(.2911)	(.2659)	(.4656)	(.4903)	(.4473)	(.3596)
Max. 10 yrs school, training	.519	.5211	.4904	.4443	.5201	.4558
	(.4996)	(.4996)	(.5)	(.497)	(.4997)	(.4989)
12 to 12 yrs school, training	.1705	.1771	.0821	.0628	.0939	.1166
	(.3761)	(.3818)	(.2745)	(.2427)	(.2917)	(.3215)
University degree	.2171	.2251	.1098	.0915	.1097	.2756
	(.4123)	(.4177)	(.3127)	(.2884)	(.3125)	(.4476)
Married	.4909	.5076	.2692	.3673	.2041	.3357
	(.4999)	(.4999)	(.4436)	(.4822)	(.4031)	(.4731)
Single	.4353	.417	.6788	.5698	.7479	.629
	(.4958)	(.4931)	(.467)	(.4952)	(.4343)	(.4839)
Divorced, widowed	.0593	.0607	.0415	.0483	.0399	.025
	(.2363)	(.2387)	(.1995)	(.2144)	(.1959)	(.1564)
No. persons in HH	3.1932	3.1435	3.8526	4.3498	3.6157	2.9576
	(1.3377)	(1.2935)	(1.6936)	(1.9007)	(1.4766)	(1.5012)
No. children in HH	.8074	.8004	.9009	1.2434	.7076	.6678
	(.9948)	(.9846)	(1.1174)	(1.2077)	(.9946)	(1.162)
Observations	77,253	71,845	5,408	2,021	2,982	283

Table 12 reports the average values of selected variables of interest for men. SG refers to the entire second generation, Turkish refers to the second generation from Turkish backgrounds, GW refers to second generation stemming from the guest-worker generation (excluding Turks), and NGW refers to the second generation stemming from all other countries.

Table 13: Descriptive statistics for women

	All	Germans	SG	Turkish	GW	NGW
Hourly wage	8.6906	8.7815	6.9621	6.6846	7.0228	8.4603
	(9.0496)	(9.1624)	(6.3023)	(5.426)	(6.7116)	(6.2551)
Gross labor income	1428.367	1442.924	1135.654	1069.668	1135.569	1619.312
	(1204.974)	(1220.574)	(775.5083)	(626.1164)	(790.5857)	(1212.313)
Unemployed (reg.)	.0752	.0754	.072	.0807	.0697	.0481
	(.2637)	(.264)	(.2585)	(.2724)	(.2547)	(.2146)
Not working	.3214	.3135	.4486	.5351	.3974	.3947
	(.467)	(.4639)	(.4974)	(.4989)	(.4895)	(.4901)
Months unemployed	.7077	.7123	.6347	.725	.6145	.3
	(2.4211)	(2.4311)	(2.2569)	(2.3846)	(2.2366)	(1.5152)
Time with firm	3.7177	3.8715	1.2349	.6344	1.5687	2.1368
	(7.1009)	(7.2028)	(4.5104)	(4.3313)	(4.5554)	(5.148)
Work hours (week)	34.3088	34.2479	35.5482	35.4093	35.4634	37.6454
	(12.5735)	(12.6457)	(10.9305)	(10.4997)	(10.8997)	(13.4221)
Age	33.4407	33.9718	24.866	24.5992	24.7458	29.3474
	(9.5753)	(9.4882)	(6.357)	(6.0024)	(6.1717)	(9.4362)
Age-group 16 to 20	.0968	.0841	.3014	.3026	.3058	.2
	(.2957)	(.2776)	(.4589)	(.4595)	(.4608)	(.4011)
Age-group 21 to 30	.3121	.2995	.5155	.5315	.5143	.4105
	(.4633)	(.458)	(.4998)	(.4992)	(.4999)	(.4932)
Age-group 31 to 40	.3362	.3469	.1635	.1528	.1677	.2211
	(.4724)	(.476)	(.3699)	(.3599)	(.3737)	(.4161)
Age-group 31 to 40	.3362	.3469	.1635	.1528	.1677	.2211
	(.4724)	(.476)	(.3699)	(.3599)	(.3737)	(.4161)
Age-group 41 to 50	.2134	.2256	.0161	.0108	.0093	.1474
	(.4097)	(.418)	(.1258)	(.1033)	(.0959)	(.3554)
Age-group 51 to 60	.0415	.0438	.0035	.0024	.003	.0211
	(.1994)	(.2047)	(.0588)	(.0489)	(.0544)	(.1439)
Max. 10 yrs school, no training	.1049	.0897	.3494	.4608	.3006	.1474
	(.3064)	(.2858)	(.4768)	(.4986)	(.4586)	(.3554)
Max. 10 yrs school, training	.5473	.5525	.4625	.4104	.4935	.4579
	(.4978)	(.4972)	(.4986)	(.4921)	(.5001)	(.4995)
12 to 12 yrs school, training	.1555	.1587	.1049	.0599	.1284	.1842
	(.3624)	(.3654)	(.3064)	(.2374)	(.3346)	(.3887)
University degree	.1923	.1991	.0832	.0689	.0776	.2105
	(.3941)	(.3993)	(.2762)	(.2534)	(.2675)	(.4088)
Married	.5657	.5769	.384	.4742	.331	.4105
	(.4957)	(.4941)	(.4864)	(.4995)	(.4707)	(.4932)
Single	.3341	.3205	.5541	.4376	.6193	.5684
	(.4717)	(.4667)	(.4971)	(.4962)	(.4857)	(.4966)
Divorced, widowed	.0814	.0839	.0405	.0476	.0387	.0159
	(.2735)	(.2773)	(.197)	(.2131)	(.1928)	(.1253)
No. persons in HH	3.2141	3.1798	3.7683	4.0971	3.6219	2.9947
	(1.2797)	(1.2435)	(1.6691)	(1.8033)	(1.5727)	(1.2193)
No. children in HH	.919	.9126	1.0226	1.2325	.934	.6263
	(1.0072)	(1.0006)	(1.1034)	(1.2)	(1.0386)	(.8435)
Observations	78,961	74,356	4,605	1,669	2,695	190

Table 13 reports the average values of selected variables of interest for women. SG refers to the entire second generation, Turkish refers to the second generation from Turkish backgrounds, GW refers to second generation stemming from the guest-worker generation (excluding Turks), and NGW refers to the second generation stemming from all other countries.

B Decomposition Analysis

Table 14: Oaxaca-Blinder decomposition of hourly wages and unemployment probabilities

	Spanish	Turkish	Yugoslav	Italian	Greek	Non GW
Decomposition of hourly wages						
Male sample						
Difference	.2908119	.31999599	.45527934	.38239707	.36242338	-.0123815
Endowments	.36405783	.36106296	.43734282	.40349705	.39095074	-.08528638
Coefficients	-.03195804	.22194864	.11659587	.12615209	.08906411	-.16419602
Interactions	-.04128789	-.26301562	-.09865935	-.14725207	-.11759147	.2371009
Female sample						
Difference	.35963228	.3111392	.45527934	.29814939	.29165346	.08521047
Endowments	.39798959	.35949019	.43734282	.32996519	.2957302	-.03168185
Coefficients	-.02338041	.26561735	.11659587	-.02111754	-.12161017	-.04756581
Interactions	-.0149769	-.31396834	-.09865935	-.01069826	.11753343	.16445812
Decomposition of unemployment probabilities						
Male sample						
Difference	.0033808	-.06687374	-.01503909	.00413786	-.00652918	-.21341288
Endowments	.0031507	-.01789276	-.00402334	.0062325	.00508955	-.02319094
Coefficients	-.20878681	.01472495	.02999264	-.00698524	-.11032114	-.21896563
Interactions	.20901691	-.06370593	-.04100839	.00489059	.0987024	.02874369
Female sample						
Difference	.	-.0088241	-.01503909	-.00079406	-.01072238	.
Endowments	.	.00683316	-.00402334	.01521719	.01424736	.
Coefficients	.	.04541707	.02999264	.00037752	-.10803844	.
Interactions	.	-.06107433	-.04100839	-.01638876	.08306869	.

A . indicates that the decomposition could not be performed as the regressors of the two groups did not have the same support.

References

Arulampalan, W. (1999). Practitioner's corner: A note on estimated coefficients in a random effects probit models. *Oxford Bulletin of Economics and Statistics 61*(4), 0305–9049.

Aydemir, A. and A. Sweetman (2006). First and second generation immigrant educational attainment and labor market outcomes: A comparison of the United States and Canada. *IZA Discussion Paper No. 2298*.

Baltagi, B. H., U. Blien, and K. Wolf (2007). Phillips curve or wage curve? Evidence from West Germany: 1980 - 2004. *IAB Discussion Paper 14/2007*.

Bauer, T. K. and M. Sinning (2008a). Blinder-oaxaca decomposition for tobit models. *Applied Economics Forthcoming*.

Bauer, T. K. and M. Sinning (2008b). An extension of the blinder-oaxaca decomposition to non-linear models. *Advances in Statistical Analysis 92*, 197–206.

Becker, G. and N. Tomes (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy 87*(6), 1153–89.

Becker, G. S. (1964). *Human Capital: A Theoretical and Empirical Analysis*. University of Chicago Press.

Becker, G. S. and B. S. Chiswick (1966). Education and the distribution of resources. *American Economic Review 56*, 358–369.

Behrman, J. R. and P. Taubman (1976). Intergenerational transmission of income and wealth. *American Economic Review 66*(2), 436–40.

Behtoui, A. (2004). Unequal opportunities for young people with immigrant backgrounds in the Swedish labour market. *Labour 18*(4), 633–660.

Blanchflower, D. G. and A. J. Oswald (2005). The wage curve reloaded. *IZA Discussion Paper 1665*.

Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources 8*(4), 436–455.

Borjas, G. J. (1992). Ethnic capital and intergenerational mobility. *The Quarterly Journal of Economics 107*(1), 123–150.

Borjas, G. J. (1993). The intergenerational mobility of immigrants. *Journal of Economic Literature 11*(1), 113–135.

Bratsberg, B., E. Barth, and O. Raaum (2006). Local unemployment and the relative wages of immigrants: evidence from the current population surveys. *The Review of Economics Studies 88*(2), 243–263.

Brenner, J. (2007). Effects of job entry restrictions on economic integration-Evidence for recent ethnic German immigrants. *Ruhr Economic Papers 25*.

Card, D. (2005). Is the new immigration really so bad? *Economic Journal 115*(507), F300–F323(1).

Card, D. and C. M. Schmidt (2003). Symposium on second-generation immigrants and the transition to ethnic minorities. *Journal of Population Economics 16*(3), 707–710.

Chamberlain, G. (1980). Analysis of covariance with qualitative data. *The Review of Economics Studies 47*(1), 225–238.

Clark, K. B. and L. H. Summers (1979). Labour market dynamics and unemployment: a reconsideration. *Brookings Papers on Economic Activity 1*, 13–60.

Constant, A. and D. S. Massey (2005). Labour market segmentation and the earnings of German guestworkers. *Population Research and Policy Review 24*(5), 489–512.

Daymont, T. and P. Andrisani (1984). Job preferences, college major, and the gender gap in earnings. *Journal of Human Resources 19*.

Drukker, D. (2003). Testing for serial correlation in linear panel-data models. *Stata Journal 3*, 168–177.

Dustmann, C. (1994). Speaking fluency, writing fluency and earnings of migrants. *Journal of Population Economics 7*, 135–156.

Dustmann, C. and A. Van Soest (2001). Language fluency and earnings equation with misclassified language indicators. *Review of Economics and Statistics 83*, 663–674.

Fertig, M. and C. M. Schmidt (2001). First- and second-generation migrants in Germany - what do we know and what do people think? *IZA Discussion Paper No. 286*.

Fertig, M. and S. Schurer (2007). Labour market outcomes of immigrants in Germany: The importance of heterogeneity and attrition bias. *IZA Discussion Paper 2915*.

Gang, I. N. and K. F. Zimmermann (2000). Is child like parent? Educational attainment and ethnic origin. *Journal of Human Resources 3*, 550–569.

Gestring, N., A. Janßen, A. Polat, and W. Siebel (2004). Die zweite Generation türkischer Migranten. *EINBLICKE 40*.

Haisken-DeNew, J., F. Büchel, and G. Wagner (1997). Assimilation and other determinants of school attainment in Germany: Do immigrant children perform as well as Germans? *Vierteljahreshefte zur Wirtschaftsforschung 66*(1), 169–179.

Haisken-DeNew, J. P. and M. Hahn (2006). Panelwhiz: A menu-driven Stata/SE interface for accessing panel data. *mimeo http://www.panelwhiz.eu*.

Halvorsen, R. and R. Palmquist (1980). The interpretation of dummy variables in semilogarithmic equations. *American Economic Review 70*(3), 474–475.

Haveman, R. and B. Wolfe (1995). The determinants of children's attainments: A review of methods and findings. *Journal of Economic Literature 33*(4), 1829–1878.

Jakobsen, V. and N. Smith (2003). The educational attainment of the children of Danish guest worker immigrants. *IZA Discussion Paper* (No. 749).

Jann, B. (2008). The blinder-oaxaca decomposition.

Kristen, C. (2000). Ethnic differences in educational placement: The transition from primary to secondary schooling. *Arbeitspapiere - Mannheimer Zentrum für Europäische Sozialforschung 32*.

Layard, R., S. Nickell, and R. Jackman (1991). *Unemployment: Macroeconomic Performance and the Labour Market*. Oxford: Oxford University Press.

Mincer, J. (1974). *Schooling, Experience, and Earnings*. New York: Columbia University Press.

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica 46*(1), 69–85.

Nielsen, H. S., M. Rosholm, N. Smith, and L. Husted (2003). The school-to-work transition of second generation immigrants in Denmark. *Journal of Population Economics 16*(4), 755–786.

Oaxaca, R. L. (1973). Male-female wage differentials in urban labour markets. *International Economic Review 14*(3), 693–709.

OECD (2005). The labour market integration of immigrants in Germany. *OECD Publishing http://www.oecd.org/dataoecd/62/12/35796774.pdf*.

OECD (2006). Where immigrant students succeed - A comparative review of performance and engagement in PISA 2003. *OECD Publishing 14 May 2006*.

Riphahn, R. (2003). Cohort effects in the educational attainment of second generation immigrants in Germany. An analysis of census data. *Journal of Population Economics 16*, 711–737.

Riphahn, R. T. (2005). Are there diverging time trends in the educational attainment of nationals and second generation immigrants? *Journal of Economics and Statistics 225*(3), 325–346.

Schmidt, C. M. (1992). Country of origin differences in the earnings of German immigrants. *Universität München Discussion Paper 92-29*.

Sweeney, K. (1996). Destinations of leavers from claimant unemployment. *Labour Market Trends October*, 443–452.

Thränhardt, D. (1989). Patterns of organization among different ethnic minorities. *New German Critique. Special Issue on Minorities in German Culture 46*, 10–26.

Uhlendorff, A. and K. F. Zimmermann (2006). Unemployment dynamics among migrants and natives. *IZA Discussion Paper Series 2299*.

Van Ours, J. C. and J. Veenman (2003). The educational attainment of second generation immigrants in the Netherlands. *Journal of Population Economics 16*, 739–753.

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

Wooldridge, J. M. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics 20*, 39–54.

Worbs, S. (2003). The second generation in Germany: Between school and labor market. *International Migration Review 37*(4), 1011–1038.