

Treude, Barbara

Conference Paper

The Effect of Ethnic Clustering on Migrant Integration in Germany

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik
- Session: Migration I, No. D05-V4

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Treude, Barbara (2014) : The Effect of Ethnic Clustering on Migrant Integration in Germany, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik - Session: Migration I, No. D05-V4, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/100584>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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

The Effect of Ethnic Clustering on Migrant Integration in Germany

Barbara Treude *

February 1, 2014

Abstract

Even though ethnic clustering is common, both economic theory and empirical research have not been able to provide a clear-cut answer on its effects on the integration of immigrants. In this paper, we study the effect of residential clustering on the labour market outcome of immigrants in Germany. Thereby, we use two measures for labour market integration: the employment probability and the wage levels. Our paper contributes to the existing literature twofold: First, we extend it to Germany on which hardly any research has been conducted. Second, we employ a new methodological approach which allows for the identification of smaller ethnic clusters and thus a more precise estimation of its effects. This is achieved by including neighbouring spatial units in the regression model. In order to control for the endogeneity of the location decision, we use a two-step strategy, combining a control function and an instrumental variable (IV) approach. We observe immigrant clustering in Germany, using a Moran's I analysis. In addition, the analysis suggests the importance of residential clustering for the employment probability of immigrants, and for second or higher generation migrants, a positive impact of clustering on wages.

JEL codes: *J61, J64, J31, R23*

Keywords: *ethnic enclaves, labour market integration, migrants, wage differentials*

*Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI) and Ruhr-University Bochum (RUB) (Barbara.Treude@rub.de)

1 Introduction

While residential segregation of immigrants is common, economic theory cannot give a clear-cut answer to whether it is beneficial or detrimental for their integration. Nor does the empirical literature provide a convincing answer. However, with respect to policy implications, the question is highly relevant. If enclaves could help first generation immigrants settle in the host country more quickly, for example, clustering might be supported by a country. In this paper, we analyse the effect of ethnical clustering on the integration of immigrants into the German labour market. Thereby, we use two measures for labour market integration: the employment probability and the wage levels.

Danzer and Yaman (2010) analyse the effect of ethnic concentration on language fluency in Germany. They find a negative relationship. To our best knowledge, there are currently no other studies pertaining to the effect of clustering on immigrant integration – either in form of language acquisition or of labour market outcome – for Germany. There is some literature on migration to Germany after World War II (*Aussiedler*), and clustering of these immigrants is generally found to have a positive effect on the labour market outcome (Dietz, 1999; Bauer and Zimmermann, 1997). However, since these immigrants are ethnic Germans, their integration into the German labour market is likely to differ considerably from the one of later immigration waves, such as the guest workers. Our paper therefore aims at making a second contribution when it comes to extending the literature on enclave effects in Germany.

In addition, we use a new methodological approach. While other studies, such as Borjas (2000), use a clustering index based on the population fraction in a respective city being of a certain ethnic group as an independent variable of an OLS estimation, we also include neighbouring indices and account for spatial effects. This allows for the use of data on a smaller regional scale which is gainful, as ethnic enclaves are not restricted to coincide with the spatial unit observed for each individual. This, in return, allows for a more precise estimation of the clustering effect, as it makes a differentiation of neighbourhoods within a city possible. In some towns, clustering of immigrants might take place, even though their absolute number is so small that the overall share of the town’s population would not suggest a cluster. In the traditional approach, these immigrants would be identified as not living in an ethnical cluster, whereas small-scale regional data can more precisely detect clustering in the data.

The remainder of this paper is structured as follows: Section 2 reviews the relevant literature. In section 3, the data we use is discussed and analysed for the presence of immigrant clustering in Germany. Section 4 describes the identification strategy and the empirical model, including a two-step strategy to deal with the non-random sorting of immigrants into neighbourhoods. Section 5 presents the results for the descriptive analysis and section 6 concludes and describes the outlook for

the paper.

2 Literature Review

From theory, the effect of ethnic enclaves on integration is not clear. If ethnic enclaves are seen as a network (Damm, 2006), they work based on information and norms. Individuals behave according to the norms in order to be a part of the network. At the same time, the network provides valuable information, such as information on job offers. Therefore, being part of that network as an immigrant can be gainful for the integration and help overcome discrimination otherwise faced. However, living in an enclave can also lower the incentives to invest in country-specific human capital of the host country, because it is not necessary in the enclave Lazear (1999). Therefore, individuals will have less chances of integration in the labour market or the schooling system. Even if the enclave might constitute a labour market of its own, where chances of labour market integration are independent from the host country's human capital, Borjas (2000) argues that these labour markets generally pay lower wages. Thus, *a priori* the effect of ethnic clustering on the integration of immigrants cannot be determined.

The empirical literature on the effect of ethnic enclaves on immigrants' labour market outcome exhibits mixed results as well. In particular for the US, many studies find the clustering of immigrants into residential areas to significantly hinders their labour market integration (Borjas, 2000; Pedace and Rohn, 2008; Sousa, 2013). However, for Sweden Edin *et al.* (2000) find a positive effect of clustering on wage after sorting of immigrants into enclaves is taken into account, and for Denmark, Damm (2006) finds positive effects on both employment probability and earnings, arguing that social networks are key for individual labour market outcomes. Another issues is whether enclaves only affect a specific group of immigrants. When looking at the absolute wage earned, Vigdor *et al.* (2008) find that enclaves mostly benefit only educated groups in the US. On the other hand, Warman (2006) finds that for Canada clustering negatively impacts the wage growth rate of immigrants, especially on high-skilled ones. Besides, the effect is found to be stronger for females. Tu (2010) stresses that clustering effects diminish over time and there is no relationship for immigrants to Canada 20 years after they immigrated. Overall, there seems to be no entirely convincing answer to the question whether ethnic enclaves hinder or benefit labour market integration.

One potential channel through which clustering affects integration is through the (non-)acquisition of language skills of the host country. For the US, Chiswick and Miller (2002) show that living in an area where many people speak the immigrant's mother tongue significantly decreases the immigrant's own skills in the host country's language. Besides, low language skills can also be shown to result in lower earnings. (Chiswick and Miller, 1996; Dustmann and Soest, 2001; McManus, 1990). On the

other hand, Bauer *et al.* (2005) argue that the causality might be reversed and find that the enclave size an immigrant chooses decreases with improvements in her language proficiency.

Overall, the effect of ethnic enclaves remains unclear both from a theoretical and empirical perspective. This is particularly true for Germany, on which there are hardly any studies. Our paper contributes to the existing literature twofold: First, we extend this branch of research to Germany. Second, we employ a new methodological approach which allows for the identification of smaller ethnic clusters and thus a more precise estimation of its effects.

3 Data

To investigate the effects of ethnical clustering on the labour market outcome of migrants, we employ data from the German Socioeconomic Panel (GSOEP)¹. The GSOEP is an annual representative household survey of individuals aged 16 and above which started in 1984. We use the geocoded version with individual address data. Immigrants can be identified by a number of indicators. For example one variable captures whether an individual was born in Germany (available since 1997, expect for 2012). In addition, the individual race (all waves) and the nationality (since 1997) are known. This allows identifying immigrants of the second or third generation. We take data for the years 2005 to 2012 and restrict our sample to those who are at least 16 in 2005 and younger than 66 in 2012. Based on these restrictions, we end up with 2,537 first generation immigrants and 1,251 second or higher generation immigrants.

The GSOEP data can be matched with microm² raster data which provides information on the share of immigrants and their ethnic background in each 1-km²-cell. The data are available for the years 2005 to 2010, and it is obtained by the analysis of pre- and surnames of the household's head. This allows for an efficient approximation of the share of foreigners in a small area. As the GSOEP sample is not stratified for regional presentation, a computation of the immigrants' share is very likely to be biased. By merging the microm data, a reliable share of the immigrants can be obtained for use in the regression analysis. Thus, the combined dataset enables an analysis on such a small regional scale that the GSOEP alone would not have allowed for.

Furthermore, in the microm data, there are several measures available for different ethnic groups. This is relevant, as it allows to identify the effect of living in an enclave with people from the same or a similar cultural background. Arguably, cross-groups effects are the lower the farther

¹The data used in this paper was extracted using the Add-On package PanelWhiz for Stata®PanelWhiz (<http://www.PanelWhiz.eu>) was written by Dr. John P. Haisken-DeNew (john@PanelWhiz.eu). See Hahn and Haisken-DeNew (2013) and Haisken-DeNew and Hahn (2010) for details. The PanelWhiz generated DO file to retrieve the data used here is available from me upon request. Any data or computational errors in this paper are my own.

²Microm Micromarketing-Systeme und Consult GmbH is a German marketing and consultancy agency.

the distance between the two groups' cultures. Therefore, using the general share of immigrants in the regression is likely to underestimate the true effect. Based on the microm variables, we group the observed immigrants into 11 ethnic backgrounds³, thus incorporating the possibility of cross-group effects between similar cultures (cf. e.g. Chowdhury and Pedace, 2007).

Figure 1 presents the overall distribution of immigrants in Germany. A clear clustering of immigrants into enclaves is visible. It can be seen that the density of immigrants is significantly higher in big cities and more populated areas. Besides, the share of immigrants is considerably higher in West Germany. This can partly be attributed to immigration patterns before the German reunification, in particular to the influx of the guest workers into West Germany in the 1960ies and 1970ies. However, recent immigration flows have also been directed towards West Germany (Bundesamt für Migration und Flüchtlinge, 2014), which indicates further clustering of immigrants.

4 Empirical Strategy

The clustering effect can be identified by comparing the labour market outcome of individuals of the same ethnic background who live in neighbourhoods that differ with respect to the share of migrants from that background. With small scale regional data, the area includes both the cell the individual is living in and neighbouring ones. The effect is averaged over the major ethnic groups in a country to find the country-specific effect of clustering on immigrants in general.

The analysis is conducted in two steps. First, we estimate the probability of being employed for individual i living in neighbourhood j in time t :

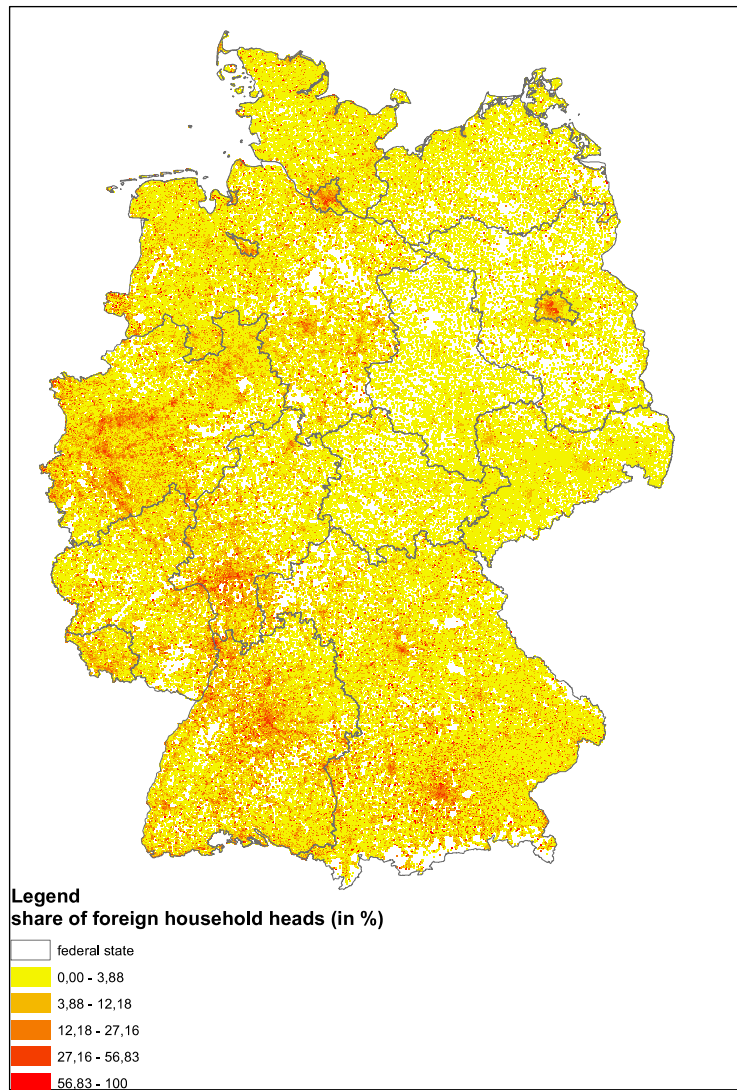
$$e_{ijt} = \alpha + C_{jt}\beta_1 + \theta_1 \sum_{k \neq j} w_{jk}^1 C_{kt} + \theta_2 \sum_{k \neq j} w_{jk}^2 e_{kt}^a + \bar{X}_{it}\beta_2 + \bar{\delta}_t + \bar{\kappa}_i + \varepsilon_{ijt}, \quad (1)$$

where e denotes employment which is one if someone is employed and zero otherwise. For the estimation we assume a logistic distribution. α is a constant, C_{jt} the clustering measure of immigrants with the same ethnic background living in j in the year t . w_{jk} is the j, k -element of the spatial weights matrix \mathbf{W} , weighting the influence of neighbourhood k on individuals in neighbourhood j . \bar{X}_{it} is a vector of controls, including age, sex, education and labour market experience. $\bar{\delta}_t$ is a vector of year dummies to account for exogenous shocks, such as overall changes of the German economy within a particular year. $\bar{\kappa}_i$ is a vector of cohort dummies, accounting for overall variation in the labour market quality of an immigration cohort (Gundel and Peters, 2007). ε is the error term.

The spatial weights matrix \mathbf{W}^1 is based on inverse distances, allowing for closer cells to have a larger effect on the individual outcome than for those less close. This seems reasonable as travel time

³The groups are: African, Asian, Balkan, Greek, Middle Eastern, Italian, Eastern European, Spanish/Portuguese, Aussiedler, Turkish, other.

Figure 1: Share of foreigners throughout Germany (2010)



generally decreases an individual's real wage, and therefore the arbitrage effect in wages decreases with the travel time. This extends to the employment probability as individuals are only willing to commute a certain distance in order to become employed. Several bounds and an unrestricted version of the weights matrix are tested and the models' fits compared. Like this, we can approximate the average spatial magnitude of cluster effects.

Arguably, apart from the interdependencies of the immigrant share, labour market outcomes also depend on those nearby. Especially on a small regional scale where individuals are sufficiently mobile, economic theory predicts that for the same individuals wages should equalise. In addition, unem-

ployed persons will commute in within area if necessary. If the regression did not account for this behaviour, the clustering effect is likely to be underestimated. Therefore, the spatial autoregressive part $\sum_{k \neq j} w_{jk}^2 e_{kt}^a$ is also included in the analysis, accounting for the influence of the average employment rate in neighbourhood k on individuals in neighbourhood j . Thereby, average employment includes the observations of both natives and immigrants. This ensures that these values do not suffer from potential non-random sorting of immigrants into specific regions. We use a different specification of the spatial weights matrix \mathbf{W}^2 based on local labour market areas (LLMA) depending on the commuting behaviour of individuals, as defined in Eckey *et al.* (2006). Thus, neighbourhoods in one LLMA are assumed to equally affect each other, yet neighbourhoods outside the respective LLMA are assumed to have no effect. This approach also controls for region-specific labour market differences.

In the second step, for those employed, we estimate the effect of the ethnical clustering on the wages, extending a Mincer’s wage equation (Mincer, 1974) as follows:

$$\ln w_{ijt} = \alpha + C_{jt}\beta_1 + \theta_1 \sum_{k \neq j} w_{jk}^1 C_{kt} + \theta_2 \sum_{k \neq j} w_{jk}^2 \ln w_{kt}^a + \bar{X}_{ijt}\beta_2 + \bar{\delta}_t + \bar{\kappa}_i + \varepsilon_{ijt}, \quad (2)$$

where $\ln w$ denotes log wages and the rest is specified as above.

4.1 Endogeneity of Location Decision

The location decision of immigrants is potentially endogenous for two reasons. First, immigrants might particularly be drawn to neighbourhoods where unobserved (labour market conditions) are particularly favourable. Thus, the error term might not be independent, resulting in an estimation bias. Second, there might be a selection of a specific group of immigrants, e.g. those that are less prone to learn the domestic language, into enclave neighbourhoods. Thus, unobserved individual characteristics may be correlated with the idiosyncratic error term, also leading to biased estimates. In order to account for both types of sorting behaviour, we plan to use a two-step approach proposed by Bayer and Ross (2006) and Bauer *et al.* (2011), combining a control function and an instrumental variable (IV) approach.

First, we try to find a measure for the neighbourhood’s unobserved characteristics to account for sorting due to these characteristics. We assume that this kind of sorting occurs on the level of the LLMA instead of the observed 1-km²-cell, as for immigrants labour market conditions are likely to be one of the key determinants. We include measures for the LLMA’s average unobserved characteristics. To this end, we run a hedonic house price regression, as shown below, where h denotes houses and l a specific LLMA:

$$\ln P_{hl} = \xi + \phi \bar{H}_{hl} + \zeta \bar{N}_l + \omega_{hl}. \quad (3)$$

We include house-specific characteristics in the vector \bar{H}_{hl} , such as number of rooms, and neighbourhood-specific characteristics in the vector \bar{N}_l , in particular the share of foreigners, average wage levels and average unemployment

We average the residuals ω_{hl} of each h over the respective LLMA l . The residuals should capture all neighbourhood-specific characteristics that cannot be observed, such as favourable labour market conditions, and are included as regressors in regressions 1 and 2. For the regression, we use data obtained from the internet portal *Immobilienscout24.de* covering almost 900,000 house offerings between 2009 and 2012. We assume that the unobserved characteristics are relatively stable over the short period between 2005 and 2007 and therefore we pool all observations observed in different years.

In order to control for individual sorting, a cell-based IV approach could be used, as proposed in Bayer and Ross (2006) and Bauer *et al.* (2011). All possible combinations of the control characteristics, such as age and education are identified, where one combination constitutes one cell. For each cell the clustering indices are averaged. These cell averages $\bar{\omega}_l$ are used in regressions 1 and 2 instead of the real clustering index. This ensures that identical individuals “live” in identical neighbourhoods, eliminating the impact of individual sorting behaviour. This kind of sorting is assumed to take place on the actual neighbourhood level, as individual preferences, such as closeness to friends, are likely to influence the location choice specific to small areas. The same cell-based IV strategy applies to the measure of unobserved neighbourhood characteristics $\bar{\omega}_l$, as individuals might have unobserved preferences for the unobserved neighbourhood characteristics.

Since the cell-base IV approach is sensitive to outliers in a relatively small sample, we also try instrumenting the share of immigrants directly. We test several instruments, including the share of old houses in the area, as done in Bauer *et al.* (2013). While this has no effect on current wages and the employment status, it is negatively correlated with the immigrants’ share, as houses in industrial centres where immigrants tend to live were destroyed during the Second World War. In addition, we test a neighbourhood’s gross domestic product (GDP) in 1900. This should also identify industrial centres in the sense of today’s immigrants’ hotspots, and it does not influence wages and employment nowadays either.

The two-step approach yields the following model for estimating the employment probability:

$$e_{ijlt} = \alpha + \hat{C}_{jlt}\beta_1 + \theta_1 \sum_{k \neq j} w_{jk}^1 C_{kt} + \theta_2 \sum_{k \neq j} w_{jk}^2 e_{klt}^a + \bar{X}_{ijlt}\beta_2 + \bar{\delta}_t + \hat{\omega}_l + \bar{\kappa}_i + \varepsilon_{ijlt}, \quad (4)$$

where \hat{C}_{jlt} and $\hat{\omega}_l$ are instrumented variables. Equation 4 is adjusted likewise. The employment probability is estimated using maximum likelihood, and the wage level regression is estimated using two-stages least squares.

4.2 Robustness Checks

As a robustness check, the information on German language skills provided by the GSOEP could be used. Thus, language skills could be used as the dependent variable to analyse the impact of ethnic clustering on skills of the host country's language. This could possibly confirm one channel through which clustering hinders labour market integration. Furthermore, the impact of clustering might be more pronounced for some demographic subgroups, such as low-skilled workers or newly arrived immigrants. The analysis will be performed with different demographic subsamples to account for such possible differences.

5 Results

The hypothesis of ethnic clustering in Germany, as expected from figure 1, is supported by a Moran's I analysis (Moran, 1950) of the share of foreigners in 2010. The analysis was conducted using a spatial weights matrix based on inverse distances, as described in section 4⁴. The significant Moran's I of 0.196 shows the clustering of immigrants clearly. Figure 2 shows the Moran scatter plot. It can also be seen visually that high shares of immigrants are likely to be surrounded by high shares and vice versa. Overall, immigrant enclaves seem to be common in Germany.

To analyse the effect of clustering on labour market integration, a descriptive analysis is conducted. The analysis first investigates employment probability and second wage levels. It already indicates the importance of clustering for German immigrants when it comes to the employment probability. For wages, the effect as shown in the descriptive analysis is less pronounced and hints at a positive impact for second generation migrants.

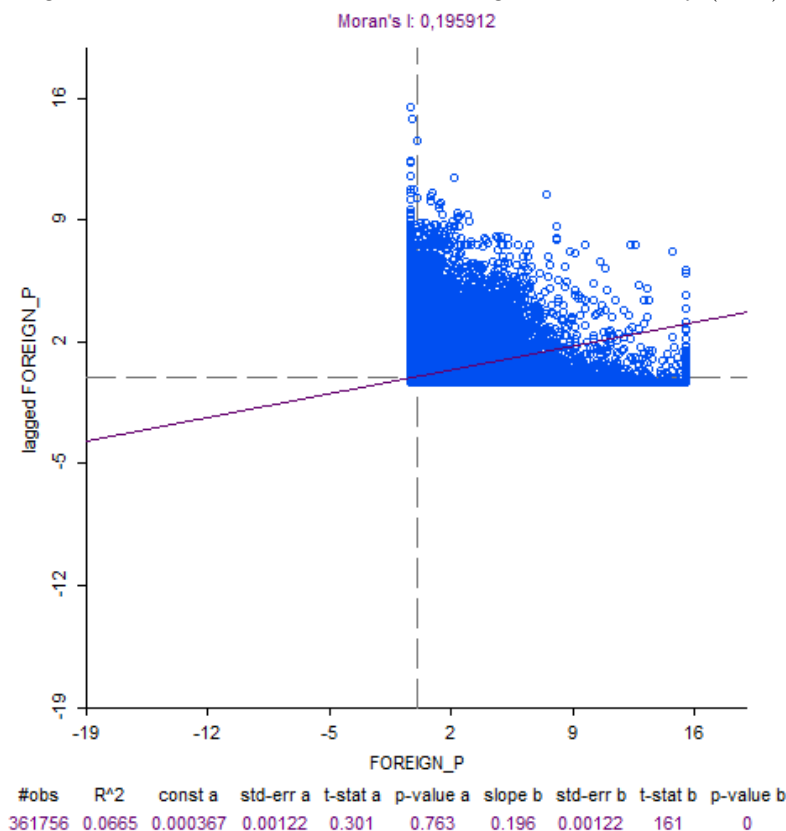
5.1 Employment Probability

Table 1 presents the summary statistics for the first regression on employment probability as described in equation 1. We observe 2,537 immigrants with a direct migration experience, i.e. first generation immigrants, and 1,251 second generation or higher immigrants. Of the first generation immigrants, most migrated to Germany several or even many years ago. On average, they have lived in Germany for 20 years when first observed.

If the first generation immigrants are compared to natives in the GSOEP, their unemployment probability is significantly higher. While it is only 10 percent for natives, it is 14 percent for first generation immigrants. This is true even though the average age (and its standard deviation) is

⁴The analysis was conducted using the software GeoDa. See Anselin *et al.* (2006) for further information. For computational reasons, a threshold was introduced.

Figure 2: Moran's I for the share of foreigners in Germany (2010)



almost the same as well as the share of women. In addition, the average years of education and labour market experience differ only by one year. Considering that the share of those living in West Germany is considerably higher among first generation immigrants (98 percent compared to 79 percent), it is likely that if accounted for the better labour market conditions in West Germany the difference in the employment probability is even higher. This hints at other factors influencing the employment probability of immigrants. We argue that one main factor is that many immigrants live in ethnic enclaves. This effect might even extend to second or higher generation immigrants, as their unemployment probability is also higher (12 percent).

5.2 Wages

Table 2 presents the summary statistics for the sample used when estimating the regression on wage levels as defined in equation 2. We observe 1,678 first generation immigrants and 865 second or higher generation immigrants. On average, the observed first generation immigrants have lived in Germany for 20 years when first observed.

Table 1: Summary statistics of the employment regression sample

Variable	Mean	Std. Dev.	Min.	Max.	N
Natives					
unempl	0.1	0.3	0	1	17053
age	40.74	11.5	17	60	17053
female	0.5	0.5	0	1	17053
west	0.79	0.41	0	1	17053
edu	12.34	2.57	7	18	17053
labexp_FT	13.88	11.45	0	45.4	17053
labexp_PT	2.4	5.10	0	45.2	17053
Direct Migration Experience					
unempl	0.14	0.34	0	1	2602
age	40.26	11.04	17	60	2602
female	0.55	0.5	0	1	2602
west	0.98	0.15	0	1	2602
edu	11.18	2.41	7	18	2602
labexp_FT	12.31	11.15	0	45.6	2602
labexp_PT	1.99	3.99	0	33	2602
ysm	19.57	10.76	1	58	2537
Indirect Migration Experience					
unempl	0.12	0.32	0	1	1251
age	34.49	11.71	17	60	1251
female	0.52	0.5	0	1	1251
west	0.93	0.26	0	1	1251
edu	11.92	2.63	7	18	1251
labexp_FT	9.13	9.98	0	45.7	1251
labexp_PT	1.87	4.54	0	36.4	1251

It can be seen that the average monthly gross wage of both first generation and further generation immigrants is lower than that of natives. First generation immigrants are comparable to natives in age and their share of women. However, their education level is also lower on average (11.35 years compared to 12.55 years). Probably their lower average wage is partly due to their lower education,

Table 2: Summary statistics of the wage regression sample

Variable	Mean	Std. Dev.	Min.	Max.	N
Natives					
wage	2476.21	1939.27	18	70000	12899
age	41.18	10.92	17	60	12899
female	0.46	0.5	0	1	12899
west	0.81	0.39	0	1	12899
edu	12.56	2.62	7	18	12899
labexp_FT	15.36	11.4	0	45.4	12899
labexp_PT	2.52	5.29	0	45.2	12899
Direct Migration Experience					
wage	2104.63	1871.16	20	42000	1722
age	40.64	10.51	18	60	1722
female	0.49	0.5	0	1	1722
west	0.99	0.11	0	1	1722
edu	11.35	2.37	7	18	1722
labexp_FT	14.03	11.16	0	45.1	1722
labexp_PT	2.3	4.37	0	33	1722
ysm	20.33	10.63	1	57	1678
Indirect Migration Experience					
wage	2269.18	1611.46	93	20000	865
age	36.21	11.41	17	60	865
female	0.48	0.5	0	1	865
west	0.92	0.27	0	1	865
edu	12.22	2.76	7	18	865
labexp_FT	10.97	10.22	0	45.7	865
labexp_PT	2.28	5.18	0	36.4	865

not indicating a clustering effect. On the other hand, as the first generation immigrants in the sample have already been in Germany for a long period, it is likely that we cannot observe the clustering effect clearly in the descriptive analysis. Channels through which the clustering effect works, such as language acquisition, tend to get less over time. Including YSM in the regression analysis, however,

we can control for this. Based on the estimated coefficient of ethnic enclaves, we may confirm one of the two hypotheses: First, enclaves only affect the employment probability, or second, enclaves only affect the labour market outcomes of newly arrived immigrants. We cannot distinguish these effects in the descriptive analysis.

Interestingly, for second or higher generation immigrants, the wage levels are closer to that of natives although natives are on average notably older (41.18 compared to 36.21) and possess longer labour market experience (15.36 compared to 10.97). Since second or higher generation immigrants also tend to live in ethnic enclaves, this may hint at a positive effect of clustering on labour market integration for this group. The effect could be driven by network effects in the enclave. We can identify this effect in the regression analysis when only considering the subsample of second generation migrants.

6 Conclusion & Outlook

Even though ethnic clustering is common, both economic theory and empirical research have not been able to provide a clear-cut answer on its effects on the integration of immigrants. In this paper, we study the effect of residential clustering on the labour market outcome of immigrants in Germany. Thereby, we use two measures for labour market integration: the employment probability and the wage level. Our paper contributes to the existing literature twofold: First, we extend it to Germany on which hardly any research has been conducted. Second, we employ a new methodological approach which allows for the identification of smaller ethnic clusters and thus a more precise estimation of its effects. This is achieved by including neighbouring spatial units in the regression model.

The Moran's I analysis confirms immigrant clustering in Germany. In addition, we observe a considerable difference between the employment probability of natives and immigrants hinting at the importance of residential clustering. With respect to wage levels, we cannot see effects for first generation migrants. This is mainly due to the fact that we cannot control for YSM and the resulting assimilation process in a descriptive analysis. We expect to find significant effects in the regression analysis. For second or higher generation migrants, the analysis suggested a potential positive impact of clustering on wages.

Besides the first descriptives, we are going to estimate the effect of clustering within one ethnic background on the employment probability and log wages. We will differentiate between men and women, as it seems reasonable to assume that both are affected differently by living in an enclave. That seems particularly true when looking at the employment probability. In order to account for the endogeneity of the location decision, we will employ a two-step approach proposed by Bayer and Ross (2006) and Bauer *et al.* (2011), combining a control function and an instrumental variable

(IV) approach. This enables us to control for sorting due to unobserved regional characteristics and unobserved individual characteristics. By testing several bounds and an unrestricted version of the spatial weights matrix based on inverse distances, we will approximate the average spatial magnitude of clustering effects.

Furthermore, we want to analyse the differences between first generation and second generation migrants. Arguably, the clustering effects are not going to be the same for first generation immigrants and subsequent generations, as the latter will acquire country-specific knowledge differently, especially so by the host country's education system. Their incentive structure is therefore possibly different. We plan to investigate both subsamples individually as well as a combined sample.

As a robustness check, we plan to use German language skills as the dependent variable and can thus possibly identify one channel through which clustering impacts labour market outcomes. In addition, we want to differentiate between several demographic groups, such as persons with different educational attainments.

References

- ANSELIN, L., SYABRI, I. and KHO, Y. (2006). Geoda: An introduction to spatial data analysis. *Geographical Analysis*, **38** (1), 5–22.
- BAUER, T. K., EPSTEIN, G. S. and GANG, I. N. (2005). Enclaves, language, and the location choice of migrants. *Journal of Population Economics*, **18** (4), 649–662.
- , FERTIG, M. and VORELL, M. (2011). *Neighborhood Effects and Individual Unemployment*. IZA Discussion Paper 6040, Institute for the Study of Labor (IZA).
- , FLAKE, R. and SINNING, M. G. (2013). Labor market effects of immigration: Evidence from neighborhood data. *Review of International Economics*, **21** (2), 370–385.
- and ZIMMERMANN, K. F. (1997). Unemployment and wages of ethnic germans. *The Quarterly Review of Economics and Finance*, **37**, **Supplement 1**, 361–377.
- BAYER, P. and ROSS, S. L. (2006). *Identifying Individual and Group Effects in the Presence of Sorting: A Neighborhood Effects Application*. Working Paper 12211, National Bureau of Economic Research.
- BORJAS, G. J. (2000). Ethnic enclaves and assimilation. *Swedish Economic Policy Review*, **7** (2), 89–122.
- BUNDESAMT FÜR MIGRATION UND FLÜCHTLINGE (2014). *Migrationsbericht 2012*. , Bundesamt für Migration und Flüchtlinge.
- CHISWICK, B. R. and MILLER, P. W. (1996). Ethnic networks and language proficiency among immigrants. *Journal of Population Economics*, **9** (1), 19–35.
- and — (2002). *Do Enclaves Matter in Immigrant Adjustment?* IZA Discussion Paper 449, Institute for the Study of Labor (IZA).
- CHOWDHURY, M. and PEDACE, R. (2007). Ethnic enclaves and labor markets: An analysis of immigrant outcomes in california. *Contemporary Economic Policy*, **25** (2), 238–249.
- DAMM, A. P. (2006). *Ethnic Enclaves and Immigrant Labour Market Outcomes: Quasi-Experimental Evidence*. Working Paper 06-4, University of Aarhus, Aarhus School of Business, Department of Economics.
- DANZER, A. M. and YAMAN, F. (2010). *Ethnic Concentration and Language Fluency of Immigrants in Germany*. IZA Discussion Paper 4742, Institute for the Study of Labor (IZA).

- DIETZ, B. (1999). *Ethnic German Immigration from Eastern Europe and the former Soviet Union to Germany: the Effects of Migrant Networks*. IZA Discussion Paper 68, Institute for the Study of Labor (IZA).
- DUSTMANN, C. and SOEST, A. v. (2001). Language fluency and earnings: Estimation with misclassified language indicators. *The Review of Economics and Statistics*, **83** (4), 663–674.
- ECKEY, H.-F., KOSFELD, R. and TÄIJRCK, M. (2006). Abgrenzung deutscher arbeitsmarktregionen.
- EDIN, P.-A., FREDRIKSSON, P. and ÅSLUND, O. (2000). *Ethnic enclaves and the economic success of immigrants - evidence from a natural experiment*. Working Paper Series 2000:9, IFAU - Institute for Evaluation of Labour Market and Education Policy.
- GUNDEL, S. and PETERS, H. (2007). *Assimilation and Cohort Effects for German Immigrants*. SOEP-papers on Multidisciplinary Panel Data Research 64, DIW Berlin, The German Socio-Economic Panel (SOEP).
- HAHN, M. H. and HAIKEN-DENEW, J. P. (2013). Panelwhiz and the australian longitudinal data infrastructure in economics. *Australian Economic Review*, **46** (3), 379–386.
- HAIKEN-DENEW, J. P. and HAHN, M. H. (2010). Panelwhiz: Efficient data extraction of complex panel data sets - an example using the german soep. *Journal of Applied Social Science Studies*, **130** (4), 643–654.
- LAZEAR, E. P. (1999). Culture and language. *Journal of Political Economy*, **107** (S6).
- MCMANUS, W. S. (1990). Labor market effects of language enclaves: Hispanic men in the united states. *Journal of Human Resources*, **25** (2), 228–252.
- MINCER, J. A. (1974). *Schooling, Experience, and Earnings*. New York, London: Columbia University Press.
- MORAN, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika*, **37** (1/2), 17–23.
- PEDACE, R. and ROHN, S. (2008). *A Warm Embrace or the Cold Shoulder: Wage and Employment Outcomes in Ethnic Enclaves*. Working Paper 08-09, Center for Economic Studies, U.S. Census Bureau.
- SOUSA, L. D. (2013). *Human Capital Traps? Enclave Effects Using Linked Employer-Household Data*. Working Paper 13-29, Center for Economic Studies, U.S. Census Bureau.
- TU, J. (2010). *The Effect of Enclave Residence on the Labour Force Activities of Immigrants in Canada*. IZA Discussion Paper 4744, Institute for the Study of Labor (IZA).

VIGDOR, J., GLAESER, E. and CUTLER, D. (2008). *When Are Ghettos Bad? Lessons from Immigrant Segregation In the United States*. Scholarly Articles 2666726, Harvard University Department of Economics.

WARMAN, C. (2006). *Ethnic Enclaves and Immigrant Earnings Growth*. Working Paper 1261, Queen's University, Department of Economics.