

Ethnic spatial dispersion and immigrant identity*

Amelie F. Constant ^{a,e,f,h}, Simone Schüller ^{b,c,d,e,h}, Klaus F. Zimmermann ^{f,g,h,i*}

^a *University of Pennsylvania, Philadelphia, United States*

^b *German Youth Institute (DJI), Munich, Germany*

^c *FBK-IRVAPP, Trento, Italy*

^d *IZA, Bonn, Germany*

^e *CESifo, Munich (Germany)*

^f *UNU-MERIT, Maastricht, The Netherlands*

^g *Maastricht University, Maastricht, The Netherlands*

^h *Global Labor Organization (GLO)*

ⁱ *CEPR, London, UK*

May 16, 2023

Abstract. The role of ethnic clustering in ethnic identity formation has remained unexplored, mainly due to missing detailed data. This study closes the knowledge gap for Germany by employing a unique combination of datasets, the survey data from the German Socio-Economic Panel and disaggregated information at low geographical levels from the last two but still unexploited full German censuses, 1970 and 1987. Utilizing the exogenous placement of immigrants during the recruitment era in the 1960s and 1970s we find that local co-ethnic concentration affects immigrants' ethnic identity. While residential ethnic clustering strengthens immigrants' retention of an affiliation with their origin (minority identity), it weakens identification with the host society (majority identity). The effects are nonlinear and become significant only at relatively high levels of co-ethnic concentration for the minority identity and at very low levels of local concentration for the majority identity. The findings are robust to an instrumental variable approach.

JEL Codes: J15; R23; Z10

Keywords: Ethnic minorities; residential segregation; ethnic identity; spatial dispersion; ethnic clustering; ethnic enclaves

Disclosure statement: The authors report there are no competing interests to declare.

* Emails: aconstan@sas.upenn.edu (Amelie F. Constant, ORCID: 0000-0002-5817-8579, LinkedIn: [ameliefotiniconstant](#)); schueller@dji.de (Simone Schüller, ORCID: 0000-0002-9842-6258, LinkedIn: [simone-schüller-6a0a58a8](#)); klaus.f.zimmermann@gmail.com (Klaus F. Zimmermann, corresponding author; ORCID: 0000-0003-2804-5692; Twitter: [@kfzimmermann](#); LinkedIn: [klauszimmermann](#));). Financial support from the German Science Foundation (DFG) is gratefully acknowledged. We thank Editor-in-Chief Wenxuan Hou, Victoria Finn, Martin Fischer, Corrado Giulietti and conference participants for many helpful comments and suggestions on previous drafts. The German 1970 and 1987 censuses used in this study were made available by the Federal Statistical Office of Germany (Destatis). We are grateful for excellent support by the statistical offices in Düsseldorf and Bad Ems. This is a substantially revised version of IZA Discussion Paper No. 7052.

1. Introduction

Western European societies as well as the U.S. and Canada are confronted with increasing diversity at an ethnic, cultural, linguistic, and religious level due to growing immigrant populations. This rise in diversity may generate economic advantages, ultimately increasing the creativity and dynamism of society. On the other hand, concerns about the existence of ‘parallel societies’ often occupy the public discourse and the media. Potential threats to social cohesion, arising from failed integration, are sometimes linked to the dense ethnic concentration of immigrants. Such concerns are equally relevant in China although Chinese migration is mostly internal. While China is largely considered ethnically homogenous, Chinese local communities differ with regards to dialect, culture and social identity, and internal migrants and foreign-born immigrants may to some extent share similar patterns (Cai and Zimmermann, 2023). Due to a long history of emigration, the Chinese diaspora is present globally, often observed in clusters or even enclaves (Constant and Zimmermann, 2016).

Ethnic clustering offers both benefits and penalties, which have been discussed in the broad literature on diasporas (Constant and Zimmermann, 2016). One argument has to do with economies of scale. Ethnic clusters may provide a sheltered environment, thus reducing the costs of economic and cultural assimilation to the host society, when migrants adjust in groups and not as individuals (Hatton and Leigh, 2011; Aydemir, 2012). Clusters also generate positive effects of communal efficacy. Their members enjoy the benefits of a reliable social network, group solidarity, and information sharing which may lead to monetary benefits. The power of social capital and friendship networks is proven to be a foundation of subjective wellbeing (Layard, 2005) and individual health (Averett et al., 2014) that are elements of sustainable development. The ubiquitous “China Towns” in every major city in the U.S. that are now offering *couleur locale* and attract tourism are examples of ethnic clustering that are vital centers of culture and commerce. They support the restaurant and real estate industries as well as they support imports and exports with China (Constant and Zimmermann, 2016).

On the other hand, research on segregation and conflict suggests that geographic ethnic clustering might offer stronger opportunities to mobilize for conflict in segregated environments since it increases the likelihood of stable coalitions (Corvalan and Vargas, 2015; Matuszeski and Schneider, 2006). Ethnic enclaves can potentially affect individuals negatively as well. For instance, newcomers may be taken advantage by the older more experienced immigrants as part of the rite of passage and can be precluded from leaving the

enclave. Han and Paik (2017) offer some more nuanced results in their study about ethnic minority concentration and economic development in China. In the western provinces, where the majority of China's ethnic minorities resides, the counties in the non-autonomous provinces had a positive and systematic correlation between changes in ethnic minority demographic concentration and changes in development compared to the ethnic minority autonomous regions. The latter have benefited less from China's "Western Development Program" as they have been predominantly inhabited by ethnic minorities and remained less integrated with the rest of China.

The empirical economic literature to date has devoted considerable efforts to identifying the impact of residential ethnic concentration on immigrant economic integration (see Chakraborty and Schüller, 2022, for an overview). Investigating alternative outcomes and exploring cultural and other dimensions of ethnic segregation, the empirical literature provides evidence of enclave effects on welfare participation (Bertrand et al., 2000), of the link between immigrant concentration and crime (Bell and Machin, 2012), of social interaction (Danzer and Yaman, 2013), educational attainment of immigrant children (Åslund et al., 2011; Chakraborty et al., 2019; Danzer et al., 2022), and ethnic occupational segregation (Zwysen and Demireva, 2020; Xu and Zhang, 2022; Zhang and Xu, 2023). Another strand of literature concerned with determinants and dynamics of ethnic clustering shows e.g. that marriages, inter-ethnic partnerships, and leaving parental home weaken ethnic clusters (Winke, 2018). Superdiversity may reduce ethnic clustering when the children of immigrants integrate into a large amalgam of ethnic groups (Crul, 2016). While immigrants are typically moving into large cities and highly urbanized areas, they are not attracted by agglomeration per se, but rather follow ethnic networks or educational and family contexts (Heider et al., 2020).

Yet, concerning the association between immigrant ethnic concentrated areas and the ethnic identity of their residents, there is still a gap in the economic literature. Notable exceptions are studies that relate immigrant concentration to the happiness of natives in Germany (Akay et al., 2014 and Akay et al., 2017). These studies show that increasing immigration rates in a locality increase the subjective well-being of natives, and that this effect is nonlinear and depends on the immigrants' ethnic identity. Another exception is Mägi et al. (2020) documenting for the Estonian context that Russian-speaking ethnic minorities in majority-dominated neighborhoods tend to more likely adjust their ethnic identity over time.

Given the economic and environmental consequences of the cultural integration of

immigrants, it is important to understand the mechanisms behind immigrants' identity formation, since immigrants are a permanent resident population in the host country. The response of immigrants – as individuals within their communities or as groups – to the process of technological development and institutional changes is inextricably related to capacity building and living sustainably. We argue that ethnic identity and cultural diversity are components of a sustainable socio-economic and environmental development. The more we understand the process of the ethnic identity and the environmental influences it endures, the better we can cater to the needs of immigrants, foster socio-economic integration, achieve full participation, and keep all members of society productive, happy, and healthy as well as the closer we get to achieving a sustainable socio-economic development.

Specifically, immigrants' ethnic identity manifests when there is a clash of cultures. Ethnic identity is how people identify in relation to “others” who may serve as reference groups. How immigrants of different ethnicities and cultures feel about the new host country and the old home country can have important ramifications on their economic, social, and political behavior. Their ethnic identity lies at the very core of their integration in the new environment as they grapple with issues.

Our motivation to study the effects of ethnic clustering on immigrant identity formation clearly relates to the wider literature on ethnic identity. The economic literature has studied the identity formation of immigrants in several countries (Constant et al., 2009, for Germany; Constant and Zimmermann, 2013, for Germany; Bisin et al., 2011, providing a rich theoretical framework), the impact of ethnic identity on labor force participation and earnings (Constant and Zimmermann, 2009, for Germany; Battu and Zenou, 2010, for the United Kingdom; Piracha et al., 2023, for Australia; Carillo et al., 2023, for Italy; Cai and Zimmermann, 2023, for China), job search behavior and occupational prestige (Pendakur and Pendakur, 2005, for Canada), and homeownership (Constant et al., 2009, for Germany). These studies show that ethnic identity has an impact on the economic outcomes of immigrants. While those totally marginalized and completely ethnically clinging to the origin do not fare well, those totally assimilated are not necessarily better off. Those with oppositional identities suffer even more penalties. Having a balance between the ethnic identity and the identity of the new country offers many more advantages.

General literature documents that ethnic identity has a beneficial impact on a variety of socio-cultural, educational, health, and psychological aspects. Chavous et al. (2003) found that African-American high-school students with a strong ethnic identity and group pride

who were also cognizant of societal discrimination had a positive academic stance and exhibited higher academic achievements than comparable youth who hid their ethnic identities, did not take pride in their group and were not aware of racial discrimination. Similarly, Altschul et al. (2006) showed that African-American middle-school students who expressed a Black identity, linked to a value for success, performed better academically than those who had a weak identity with their ethnic group and who did not perceive academic achievement as related to their ethnic group identity. Lastly, studying youth in immigrant families from Asia and Latin America in the U.S., Fuligni (2011) found that those with a strong ethnic identity exhibited positive psychological outcomes such as high self-esteem, a greater sense of happiness, and better academic adjustment.

Our study examines the impact of the local co-ethnic concentration on the formation of immigrants' feelings of belonging to the host society and culture as well as on the retention of their affiliation to the origin culture, and any other combination of their ethnic identity. It contributes to literature in several ways. First, by comprehensively documenting the extent and pattern of ethnic residential clustering among guestworker immigrant groups in pre-unification West Germany we offer unique insights into the behavior of labor immigrants who live in ethnic clusters and how they engage and interact with the host society. Second, by exploiting unique data, we are able to use information not only on cities or employees but also on the *total* resident population in West Germany, which affords us broader external validity. Third, we are able to examine the link between ethnic clustering and immigrants' ethnic self-identification at the more disaggregated level of counties and investigate nonlinearities in this relationship. Thus, we provide new knowledge on the topic.

A serious methodological hurdle in this type of research is that immigrants' locational choices may be influenced by networks, job opportunities, and other factors. Consequently, empirical studies suffer from endogeneity bias. In our study, we are able to address inherent endogeneity issues by using a unique combination of datasets, namely: survey data from the German Socio-Economic Panel and unexploited data from the full German censuses of 1970 and 1987. The latter provide useful information at a disaggregated geographical level. They are also the last administered full censuses in Germany, first because the country changed after unification on 3 October 1990, and second because of the growing resistance in the German population to full censuses. Consequently, we focus on the historical experiences that West Germany has made with its mass labor immigration in the 1960s and early 1970s.

We circumvent the issue of immigrant self-selection into certain locations by exploiting

the quasi-experimental setting of the historical exogenous geographical distribution of immigrants during the recruitment phase in the 1960s and early 1970s. We argue that because recruitment was purely demand-driven and foreign workers (or guestworkers) were allocated to specific firms in specific locations before arriving in West Germany, the initial geographic placement of immigrants was exogenous to unobserved individual characteristics such as willingness to culturally integrate.

Forth, we explore the existence of potential nonlinearities in the enclave effect on ethnic identity. By nonlinearities we mean that the increase in ethnic residential concentration may reach a critical level or saturation in which the self-identification of immigrants may change. To date, little attention has been paid to the empirical analysis of nonlinearities or threshold effects in economic research concerned with ethnic segregation.

We have organized our paper as follows. In Section 2, we provide the background on the ethnic concentration and the guestworker allocation in West Germany, supported by literature. In Section 3, we describe the data, the empirical setup, and summary statistics. In Section 4, we present the method and the results and discuss robustness checks. We summarize and conclude in Section 5.

2. Ethnic Clustering and Historical Background

2.1 Ethnic clustering and the formation of ethnic identity

Bisin et al. (2011, 2016) and Bisin and Verdier (2023) incorporate several theories of identity formation from psychological and sociological research and propose a theoretical framework, which allows for two distinct motivational mechanisms of group identity formation: *cultural conformity* and *cultural distinction*. The former is based on “assimilation theories, in political science and sociology [...] and contact theory in social psychology” (see also Akerlof and Kranton, 2000, 2010 for an incorporation into a model of economic behavior) whereas the latter is based on “theories of multiculturalism [...], and conflict.” (Bisin et al., 2016, p.147). These two mechanisms provide contrasting empirical implications on how residential ethnic segregation might affect immigrants’ ethnic identity.

Under the mechanism of *cultural conformity*, a high degree of ethnic clustering is likely to strengthen in-group loyalties and motivate immigrants to retain their respective minority identity. With lower concentrations of co-ethnics in a residential area, group boundaries would increasingly blur due to social interaction with natives and progressively lead to the adoption of the host country culture. In contrast, if *cultural distinction* is the mechanism at play, residing in an area with a low co-ethnic concentration would induce a psychological

cost associated with the exposure to cultural differences. Retaining one's own distinctive cultural heritage would then help reduce this cost. Immigrants residing in an area with a relative high density of co-ethnics would have fewer incentives to preserve their minority affiliation, which might in turn favor the process of cultural integration.

Previous empirical evidence on the link between ethnic segregation and immigrant identity formation is extremely sparse and provides mixed results. Based on UK data, Bisin et al. (2011, 2016) show that immigrants' minority identity may be more intense in mixed rather than in segregated neighborhoods, thus proving the cultural distinction hypothesis. This result is consistent with non-economic literature that uses U.S. data on Latino and African Americans and finds that racial-ethnic segregation weakens racial-ethnic identity (Oyserman and Yoon, 2009). In contrast, Battu and Zenou (2010) provide evidence for cultural conformity showing that living in an ethnic enclave is associated with very low levels of British identification and a rather strong affiliation with the respective ethnic group.

Koczan (2012) examines the role of ethnic identity among Turkish and former Yugoslavian immigrants in Austria and Germany. She finds that ethnic identity has no significant effect on education, employment, and political orientation, making her conclude that a strong ethnic/religious minority identity does not necessarily act as a constraint. Putting it differently, she shows that the immigrants' socioeconomic integration is not hindered by how strongly they feel Turkish or Serbian; their political orientation is not affected either.

In her case study about Turks in Duisburg in the Ruhr area of Germany, Ehrkamp (2005) documents that while Turks have transformed their neighborhoods through the establishment of communal places, through transnational consumption and mass media into conspicuously Turkish neighborhoods, at the same time, they also engage with the German society. Going beyond the sometimes-conflicting attachments Turks have towards their home and host country, the author discovers that ethnic belonging, transnational ties and engagement with the host society are complementary rather than mutually exclusive situations.

Examining immigrant clustering throughout Germany on social integration, Danzer and Yaman (2013) were the first to exploit the quasi-experiment of guestworker recruitment. While their results suggest that a high co-ethnic concentration reduces the likelihood that immigrants will socially interact with natives, they find no significant impact of living in an ethnic enclave on individuals' self-reported ethnic identification. This somewhat puzzling

result might be related to the following facts. First, the authors employ a rather wide concept of ethnic enclaves, which is being measured at the regional level of the so-called *Raumordnungsregionen*.¹ Second, the impact might be in fact nonlinear and only become significant at relatively high levels of co-ethnic concentration or after a threshold is reached. Third, the authors use information from a 2%-sample of the German employee population, which does not allow for an accurate calculation of local ethnic composition at the geographic level of counties.

2.2 Ethnic concentration in West Germany

The public debate in many Western European countries often focuses on the integration of immigrants. One argument is that immigrant concentration in certain neighborhoods or counties can hinder familiarization with the host country and learning the language as well as it may give rise to oppositional identities and attitudes. In the German context, the fear of parallel societies and the formation of ethnic neighborhoods as bulwarks are repeated topics in the national media (see e.g. Hanke, 2023 or Mahl Zahn, 2023).

There is no clear consensus in the empirical literature about the definition of ethnic concentration, ethnic clustering or enclaves. An enclave could be a pure microcosm of a foreign country within an area of the host country or a looser gathering of several different ethnic groups. In Germany, there are no “China towns” or “little Italies” as in the U.S.; there are no big urban conglomerates as New York City either. Yet, little is known about the existence of ethnic clusters in Germany and their actual effects. Limited availability of relevant data is one culprit, as data by immigrant nationality and geographically disaggregated levels are extremely rare. The few studies that exist document a stable, but rather low degree of ethnic concentration in Germany. Therefore, the term “ethnic concentration” appears to be more adequate than “ethnic enclaves” in the German context.

Schönwälder and Söhn (2009) examine immigrants’ settlement structures in Germany based on data from the 2005 Microcensus as well as on city-level data from the Inner-city Spatial Observatory (IRB). They document low levels of ethnic residential concentration and segregation within West-German cities with most immigrants residing in mixed neighborhoods. Interestingly, the authors find the immigrants’ residential structures still reflect the labor demand pattern of the guestworker recruitment era from the 1960s and

¹ *Raumordnungsregionen* are located at the geographical aggregation level between NUTS-2 (*Regierungsbezirk*) and NUTS-3 (*Kreis – county*) regions. There are 96 *Raumordnungsregionen* in Germany, and they are constituted by grouped counties (NUTS-3). On average, they comprise approximately 500,000 inhabitants.

1970s. For example, the major settlement locations for the Turkish immigrants still appear to be the Ruhr region, Cologne, Hamburg, and Berlin, while Italians and ex-Yugoslavs are more heavily populated in Southern Germany. Furthermore, and in contrast to other European countries (or the U.S.), the authors show that the immigrant population in Germany is not predominantly concentrated in a few urban centers, but rather distributed over a large number of cities, many of them small- and medium-sized. These results are confirmed by more recent evidence based on geocoded data (Buch et al., 2021).

Glitz (2014) studies the overall levels of ethnic workplace and residential segregation among the active labor force population in Germany using administrative data at the municipality level. His results show that ethnic minority workers are less residentially segregated than they are at their workplace across establishments. These levels of segregation appear to be relatively stable over the period 1975 to 2008. Interestingly, low-educated workers are significantly more segregated than the high educated across workplaces, while this is not the case across residential locations.

2.3 The quasi-experiment of guestworker immigrant allocation in West Germany

Germany's post-second-war reconstruction period made necessary the need for extra labor. A demand-driven recruitment process, the guestworker scheme, started in the early 1950s; it became more intense after the erection of the Berlin Wall in 1961, which cut off many workers from East Germany and other eastern-bloc countries. It was operationalized by a commission of the Federal Labor Office in cooperation with the respective national labor authorities. The commission received requests for workers by German firms and then accordingly allocated applicants to specific employers. Screening and recruitment was taking place in the sending countries. Residence permits were issued for the initial duration of one year and were conditional on employment with the assigned firm in a specific locality.

Guestworker bi-national treaties were signed with Italy (in 1955), Greece and Spain (in 1960), Turkey (in 1961), and the former Yugoslavia (in 1968) to meet the increasing labor shortages in the rapidly growing post-war German economy. Initially, these blue-collar workers were housed in barracks. While eventually they could rent apartments and stay longer, they did not change geographic locations. In the meanwhile, some guestworkers started bringing in their families, who stayed with them in the same location. After the labor migration ban in 1973, many of the guestworkers, particularly the Turks, who were not EU members, stayed in Germany. Since the ban would forbid them from coming to Germany if

they returned to Turkey, they brought their families in Germany. Note that family reunification was also predated upon an existing apartment, residential registration, and a job, among other stipulations.

Therefore, the allocation of immigrants across Germany was exogenous to their volition to live in a specific area. In fact, the guestworker program was explicitly designed to accommodate temporary, not permanent, immigration, hence its name. The guestworkers themselves did not intend to stay permanently in Germany and thus were willing to accept the conditions of the move and job types (Constant, 1998).

Furthermore, internal migration in Germany has been, and still is, very low. Immigrants and native Germans do not change locations often. Immigrants in particular, after they become accustomed to a certain area and build ties with the other local co-ethnics find it harder to pack and move to a new location and go through the motions of being foreigners again. These facts argue that there has been little movement of immigrants within Germany.

3. Data and Empirical Strategy

3.1 The Data

We employ a combination of individual-level data from the German Socio-Economic Panel (SOEP) (Goebel et al., 2019) and county-level data from the full German Census. 1987 is in fact the only year, for which the available data allow for such a combination. Survey information on immigrants (including information on their ethnic identification) are available annually since 1984 in the SOEP, which is nationally representative. At the same time, geographically disaggregated administrative data on the county-level ethnic concentration of different ethnicities are available from the West-German full censuses of 1970 and 1987 only. Due to the West and East-German re-unification that occurred in 1990 and the rising resistance in the German population to detailed recording of individual private information, a full census has not administered ever since. The focus on West-Germany makes also sense due to the fact that migrants are traditionally non-existent in East Germany.

Using the 1987 full German census enables us to define ethnic concentration down to the county-level encompassing 328 West German counties (236 rural and 92 urban). This provides us with the unique opportunity to precisely calculate the exact number of immigrants at the county-level. In Germany, counties are an intermediate administrative level between federal states and municipal governments and correspond to the NUTS-3 level (similar to zip-code information). In contrast to the SOEP data, the full German census

data do not contain information about the country of birth of individuals, but only about citizenship for the four largest immigrant nationalities at the time: Turkish, ex-Yugoslav, Italian, and Greek. We treat those nationalities in the sequel as ethnicities. Note however that in 1987, the share of naturalized immigrants in Germany was very low due to the prevailing law of *ius sanguis* (nationality by descent) and other monetary and bureaucratic hurdles that were then in effect. It is therefore safe to assume that the number of foreign citizens and the number of foreigners born in a particular country of origin are very highly correlated.²

The immigrant sample of the SOEP data contains rich information on foreign-born individuals from the four aforementioned largest ethnicities in the 1987 full German census. We, therefore, focus our analysis on these four groups. Most importantly, the SOEP has information about self-assessed German and ethnic identity by all respondents, i.e., feelings of belonging to the host country society and origin country culture, for those foreign-born with a non-German passport.³ Answers are measured on a five-point scale which we use in a reversed order ranging from “no” (1) to “full” (5) identification with respect to the host country (West Germany) and the corresponding country of origin. The SOEP provides information on a larger number of other individual characteristics, which we use and explain later below.

3.2 Basic information on residential ethnic segregation in West Germany

To describe the extent of the ethnic residential segregation in West Germany and identify some basic patterns we rely on the Full German Census of 1987. According to Panel A of Table 1, 6.79% of the resident population held a foreign passport. The largest immigrant group was the Turks, who constituted 2.35% of the total population, followed by the ex-Yugoslavs (0.90%), Italians (0.82%) and Greeks (0.42%). Panel A further illustrates settlement patterns across counties. Foreign citizens represented, on average, 5.33% of the population in one county. The highest share of foreigners in one county was 20.40%. The highest population share of one ethnic group (the Turks) in only one county was 7.58%.

² Note that the group of foreign citizens in the 1987 Full German Census also includes German-born individuals who hold a foreign passport.

³ The exact questions are: (1) “To what extent do you feel German?” with the options: *I feel fully German, I feel mostly German, I feel partly German, I hardly feel German, I do not feel German at all*; and (2) “To what extent do you feel [e.g. Turkish] here in Germany?” with the options: *I feel fully [e.g. Turkish], I feel mostly [e.g. Turkish], I feel partly [e.g. Turkish], I hardly feel [e.g. Turkish], I do not feel [e.g. Turkish] at all*.

Table 1 about here

To what extent did immigrants cluster in “high co-ethnic concentration counties” in 1987 West Germany? To answer this question we apply a relative measure and define three levels of co-ethnic concentration county. A *high* co-ethnic concentration is one in which the share of a specific ethnic group is at least twice as high as in Germany as a whole.⁴ The national share of co-ethnics serves as a benchmark since it is the share that one would expect if the respective ethnic group was uniformly distributed across the country. *Medium* local co-ethnic concentration is then defined as an area in which the share of co-ethnics among the population is higher, but less than twice as high, than the national share of the respective group. A county with a *low* local concentration of co-ethnics has a co-ethnic population share lower than the national share.

Table 1 demonstrates that the majority of each immigrant group lives in either medium or high concentration counties, that is, in counties with a co-ethnic share that is higher than the national level. When it comes to high concentration areas, however, we observe some differences across ethnic groups. While for ex-Yugoslavs, Italians, and Greeks the share of living in such counties amounts to 40 – 50%, among Turkish nationals it is only 21.19%.⁵ Overall, the descriptive evidence suggests that most immigrants tend to cluster and agglomerate in areas with relatively higher shares of co-ethnic residents, but the clustering is stronger for the smaller ethnic groups (ex-Yugoslavs, Italians, and Greeks) than for the largest one (Turks), by far.

Nonetheless, the “high ethnic concentration” counties in Germany are relatively few, which is in line with previous findings (e.g., Schönwälder and Söhn, 2009; Buch et al., 2021). Buch et al., (2021) also found a negative relationship between ethnic diversity and segregation in German cities. For each of the four ethnic groups under study, Figure 1 displays the distribution of counties according to the counties’ population shares of the respective ethnic group. The histograms show that the majority of German counties have in fact low levels of ethnic clustering. The dashed vertical line in each graph represents the overall national population share of the respective ethnic group. We define counties with local ethnic population shares below this benchmark as “low co-ethnic concentration

⁴ Borjas (1998), Edin et al. (2003) and Schönwälder and Söhn (2009) employ a similar definition.

⁵ Based on a comparable measure of local ethnic concentration in terms of area zip codes, Borjas (1998) finds that 48% of US residents with a migration background lived in relatively high concentration areas in 1979 with significant dispersion across ethnic groups (e.g., 83.8% of Mexicans, 49.6% of Italians, compared to only 25.8% of Greeks). Applying the same calculation for Sweden, Edin et al. (2003) show that in 1997 42% of first-generation immigrants resided in ethnic enclaves.

counties”. With respect to all four ethnic groups, the cumulative distributions of local ethnic population shares show that at least 70% of counties qualify as “low co-ethnic concentration counties” according to our definition. The solid vertical lines in Figure 1 indicate twice the overall national population share of the respective ethnic group. We define counties with ethnic population shares below (above) this second threshold as “medium (high) co-ethnic concentration counties”. It is obvious from Figure 1 that there are many fewer counties with medium or high ethnic concentration than low concentration areas. There is, hence, suggestive evidence of ethnic clustering in Germany. Turks have a much flatter distribution of ethnic concentration across counties than the other ethnicities.

Figure 1 about here

Moreover, we observe marked differences across ethnic groups with respect to their spatial distribution across Germany. The top-five counties in terms of co-ethnic shares are displayed for each ethnic group at the bottom of Panel A in Table 1. Note that there are no megacities in Germany. In fact, in 1987, no German city had more than 2 million inhabitants and there were only 13 cities with more than 500,000 inhabitants. Out of the latter, only six are represented in the top-five counties with highest co-ethnic concentration. We find that Turks are prevalently residing in the Ruhr area, Cologne, and Berlin, whereas the other ethnic groups tend to reside in the Southern German counties, possibly reflecting regional labor demand by recruiting firms during the guestworker program. While Buch et al., (2021), in general, find a negative relationship between ethnic diversity and segregation in German cities, they detect in the old industrialized Ruhr area a low level of ethnic diversity among foreign workers with comparatively high segregation.

Differences in the spatial distribution of ethnic clusters across ethnic groups are also evident in Figure 2, which is a graphical illustration of ethnic concentration levels in counties across Germany for each of the four ethnic groups. As noted also by Schönwälder and Söhn (2009), the geographical pattern of ethnic groups largely reflects the labor demand pattern of the guestworker recruitment. For example, the first guestworker treaty with Italy recruited mainly agricultural workers for Southern regions. Note that we exploit these differential concentrations of ethnicities across counties in our analyses by using only differences in ethnic concentration levels within the same county. As observed in Figure 1, we again note here that immigrant groups tend to cluster in few high concentration counties.

Figure 2 about here

3.3 *Sample and descriptive statistics*

We base our analysis on a sample of first-generation migrants, aged 16 and older, who migrated from one of the four major guestworker source countries: Turkey, the former Yugoslavia, Italy, and Greece. We exclude individuals with missing information on the year of immigration, educational level and ethnic identity. After imposing these restrictions, the final sample based on SOEP data consists of 1,881 individual observations. Local-level information from the German full census is then matched to each of these observations based on the county of their residence. Among the 117 counties in our sample, 76 are rural and 41 urban. The average number of observations in a county is 16.09. Since the full universe of the 328 counties is not represented in the SOEP sample, Panel B of Table 1 reproduces the indicators presented in Section 3.2 for the 117 counties represented in the SOEP sample in 1987. We observe no indications of a potential bias.

We report summary statistics of the key variables by degree of local co-ethnic concentration in Table 2. As in Table 1, we define a “low degree” of co-ethnic concentration for each ethnic group as counties in which ethnic concentration – that is, the size of the ethnic group relative to the population in the respective county – is lower than the share of this ethnic group in the entire population. A “high degree” of co-ethnic concentration, on the other hand, entails counties where the local-level share of a respective ethnic group is at least twice as high as the national population share of this group. By comparing the fraction of the co-ethnic population within a county to the fraction that one would expect if the ethnic group was distributed randomly across counties, we basically construct an individual-level measure of segregation similar to Borjas (1998).

Table 2 about here

The first rows of Table 2 compare the mean levels of the self-assessed strength of identification with Germany (German identity) and the country of origin (minority identity) for immigrants residing in a low, medium or high co-ethnic concentration county. With respect to German identity, there seems to be no differences, on average, between those individuals residing in a county of low co-ethnic concentration and those living in a medium or high concentration county. Immigrants’ affiliation with the country of origin appears to

be weaker in counties with a medium or high concentration of co-ethnics as opposed to those in which co-ethnics are relatively rare. This first descriptive evidence is in line with the theoretical and empirical findings of Bisin et al. (2011, 2016) with minority identity being stronger in mixed rather than in segregated areas (*cultural distinction*).

Table 2 contains the list of a larger set of control variables and individual characteristics that we use in the econometric analysis. We follow previous literature concerned with factors of ethnic identity formation (Zimmermann et al., 2008) and include both pre- and post-migration characteristics of individuals. Pre-migration characteristics are gender, age-at-entry, and dummies capturing human capital acquired in the country of origin as well as ethnicity. Post-migration characteristics are years-since-migration (YSM), a dummy variable indicating whether an individual experienced an employment spell in Germany,⁶ and indicator variables capturing human capital acquired in Germany.

Investigating the distribution of pre- and post-migration characteristics across counties with different degrees of co-ethnic concentration offers a first indication of whether immigrants are sorted into specific locations along the lines of these observable characteristics. While there seem to be no differences in terms of age at migration or gender, interesting and diverse patterns emerge for the four ethnic groups. Turks in our sample are more likely to reside in medium concentration counties rather than in areas with low or even high co-ethnic concentration. Yet, no particular preference emerges among individuals from the former Yugoslavia. Italian as well as Greek nationals are more likely to reside in high rather than medium or low co-ethnic concentration areas. This roughly reflects the pattern observed in Table 1 (Panel A), based on data from the 1987 full census.

There is no indication of distributional differences with respect to pre-migration or post-migration education, with the exception that individuals with incomplete schooling in their respective home country are more likely to reside in areas with a high co-ethnic concentration. When it comes to post-migration characteristics, the only significant differences are in YSM with individuals who have been living in Germany for longer being more likely to reside in high concentration counties. Summarizing the above, there seems to be only weak concern about self-selection in co-ethnic clusters according to the observed characteristics. This is consistent with the descriptive observations by Schönwälder and Söhn (2009) and the findings of Danzer and Yaman (2013, 2016).

The bottom panel of Table 2 presents local characteristics by degree of ethnic

⁶ We use SOEP biography data on employment spells to generate an indicator variable that equals one if the individual had worked at any point in the period between the year of immigration and 1987.

concentration. The local native population density, namely the number of German citizens per square kilometer, strongly increases for medium and high co-ethnic concentration relative to low co-ethnic-concentration counties, which indicates that these are mostly urban areas. Therefore, it is not as surprising that the local unemployment rate is significantly lower in counties with a relatively high concentration; these are at the same time the more urbanized locations. The size of the local immigrant population increases with higher co-ethnic concentration, suggesting that ethnic clusters are likely to be immigrant clusters as well.

From the descriptive evidence, we conclude that sorting according to observed individual characteristics does not seem to be a major problem for the analysis. However, since individuals residing in low and those living in high co-ethnic concentration counties face very different local characteristics, it is important to carefully control for such potentially confounding factors at the local level. Furthermore, since we find considerable differences across ethnic groups, we control for general systematic cultural differences across groups by including ethnic group fixed-effects throughout the following empirical analysis.

4. Empirical Investigations

4.1 Identifying effects of local ethnic concentration

To estimate the relationship between local co-ethnic concentration and immigrants' subjective strength of affiliation with the host country (German identity) and the country of origin (minority identity) we analyze variation *within* counties across ethnic groups rather than variation across counties. Using within-county variation has the major advantage that enables us to control for observed as well as unobserved common factors at the local level, which affect each ethnic group in the same way. Our estimation equation is thus specified as:

$$(1) \quad ID_{iek} = \alpha * \ln(\text{ethnic group size})_{ek} + \beta * X_{iek} + \mu_k + \gamma_e + \varepsilon_{iek}$$

ID_{iek} represents the self-assessed identity measure (German or minority, respectively) of individual i of ethnicity e who lives in county k ; $\ln(\text{ethnic group size})_{ek}$ is the natural logarithm of the number of co-ethnics of ethnicity e in county k ; X_{iek} is a vector of individual characteristics; μ_k is a county fixed effect; and γ_e is a fixed-effect for each ethnic group. Unobserved determinants of the individual's ethnic identity measure(s) are captured

by the error term, ε_{iek} . In this specification, as in all proceeding analysis, we cluster the standard errors at the local (county) and ethnic-group level.

By including fixed-effects for each county (μ_k) and for each ethnic group (Y_e) at the same time, we exploit only variation in local ethnic concentrations that is not systematic across ethnicities or across counties. Hence, we control for the average systematic differences across counties and across ethnic groups in any observable or unobservable variable.

Following Edin et al. (2003) and Damm (2009), equation (1) employs the log of the local co-ethnic group size as the key explanatory variable and imposes log-linearity. However, some evidence on enclave effects (Bell and Machin, 2012) suggests the existence of potential nonlinearities. Bisin et al. (2016) find significant nonlinearities as well in the relationship between ethnic concentration and minority identity in the UK. To explore nonlinearities in this context, we employ the enclave measures defined above in Sections 3.1 and 3.2 (see also Tables 1 and 2) indicating medium and high co-ethnic concentration relative to the national population share of each ethnic group. We consider

$$(2) \quad ID_{iek} = \alpha_1 * MedC_{ek} + \alpha_2 * HighC_{ek} + \beta * X_{iek} + \mu_k + Y_e + \varepsilon_{iek},$$

where $MedC_{ek}$ is an indicator variable equal to one if the population share of ethnic group e in county k is higher, but less than twice as high, than the national population share of this group; $HighC_{ek}$ is a dummy variable indicating whether the ethnic concentration of group e in county k is at least twice as high as its national share.

In equations (1) and (2), local (co-)ethnic concentration is first considered to be exogenously determined since the guestworker recruitment scheme allocated immigrants to specific firms across Germany. Indeed, Table 2 revealed no evidence of sorting according to observable characteristics, which is consistent with our assumption of exogenous placement. This is corroborated by Danzer and Yaman (2013, 2016) and Danzer et al. (2022) who find no evidence of negative sorting in the German context of guestworker immigrants and show that changes in immigrants' locations in the years after recruitment were not selective among observable characteristics.

Nonetheless, a selection bias might still affect the estimates if, during the years after initial placement, individuals had systematically sorted into or out of enclaves according to unobserved factors related to ethnic identity formation. Additional immigration flows in the course of family reunification period after the end of the guestworker scheme could

potentially bias our results if inflows into highly concentrated areas were systematically different in terms of unobservable characteristics from inflows into counties with lower concentrations of co-ethnic residents.

To address these concerns, we employ an instrumental variable (IV) strategy as a robustness check. For such a strategy to be successful, we require a variable that significantly influences local co-ethnic concentration in 1987, but at the same time it is not directly associated with the dependent variable for other, independent reasons. We instrument the local enclave measures using data from the 1970 Full German Census. That is, we use county-level ethnic concentration shortly before the guestworker recruitment was banned in 1973 to predict local ethnic segregation levels in 1987. Note that the initial geographic distribution of guestworker immigrants has been exogenous throughout the recruitment phase and these settlement patterns have been strongly persistent since then.

By implementing this instrument, we face the practical difficulty of several reforms of county territories being executed between 1970 and 1987 in some federal states. We are able to retrace these changes in cases where entire counties or their main parts were merged and necessarily ignore relatively minor changes. In practice, we match each 1970-county to a county in 1987 and compute enclave indicators based on the 1970 information. We then use the resulting local enclave measure to instrument the respective enclave measure for each county in 1987.

4.2 Baseline estimates

Table 3 presents the main baseline estimation results according to equations (1) and (2). The analysis relates the county-level co-ethnic concentration experienced by immigrants of the four main guestworker ethnicities under study (Turkish, ex-Yugoslav, Italian, and Greek) to the strength of their subjective identification with Germany (majority identity) or with their respective culture of origin (minority identity). Columns 1 and 3 contain the results from a log-linear measure of ethnic concentration. Columns 2 and 4 include indicators of medium and high residential concentration to test for potential nonlinearities. Both types of specifications control for sets of demographic, pre- and post-migration characteristics, and include both county and country-of-origin fixed effects. Throughout the analysis, the reported standard errors are cluster-robust at the ethnic group and county level.

Overall, our results show that immigrants residing in a county in which their specific ethnic group is relatively strongly concentrated feel less German and more affiliated with their culture of origin when compared to immigrants residing in the same county, but whose

ethnic group is relatively less present in this location. We clearly find a nonlinear enclave effect with respect to both the German and minority identity. Interestingly, the effects only become significant at relatively high levels of co-ethnic concentration (with respect to the minority identity) and at very low levels of local concentration (with respect to the majority identity) respectively. That is, the nonlinear pattern is different for host and home country identity.

When we allow for nonlinearities in Column 2 of Table 3, we find that immigrants facing medium or high co-ethnic concentrations in their county of residence have, on average, significantly lower levels of German identity than those in low co-ethnic concentrations. Remarkably, there appears to be no further significant difference in German identity between the medium and the high category. Within a county, the difference in German identity between an immigrant belonging to a low-level concentrated group and an immigrant experiencing medium or high levels of local co-ethnic segregation amounts to about 45% of one standard deviation. In turn, an immigrant has a relatively strong German identity when residing in a low co-ethnic concentration county, *ceteris paribus*.

Column 4 of Table 3 shows the nonlinear pattern with respect to the minority identity. In this case, we find no significant differences between low and medium co-ethnic concentration. At very high levels of local ethnic concentration, however, an enclave effect becomes apparent and significant. In terms of magnitude, residing in such an enclave increases an immigrant's strength of a minority identity by 33.4% of a standard deviation.

Table 3 further reveals that a relatively stronger German identity is generally more prevalent among immigrants with more time in the host country, those who immigrated at younger ages and those with more education. While education acquired in both home and host countries plays a positive role, it is education acquired in Germany that has predictably stronger effects. Whether or not an immigrant had an employment spell at any time since migration plays a positive but not significant role. These patterns are mainly mirrored in the opposite direction with respect to the minority identity. A notable exception is schooling acquired in Germany, which does not play a significant role for immigrants' tendency to retain a strong affiliation with the culture of origin.

Compared to immigrants from Turkey (reference category), immigrants from the former Yugoslavia exhibit, on average, a stronger affiliation with Germany and weaker retention of ties with their country of origin. The Greeks, on the other hand, are less inclined to identify with Germany and tend to retain stronger ties to the Greek culture than Turkish immigrants do. There are no marked differences in ethnic identity between Turks and Italians.

4.3 Instrumental variable (IV) estimation

Table 4 provides the instrumental variable estimates for German and minority identity. In this exercise, we exploited the exogenous initial guestworker placement across German counties during the recruitment years (based on previous respective co-ethnic concentrations measured in the 1970 Full German Census). Since in our main analysis we found that effects are nonlinear and become significant only at very low levels of local concentration for the majority identity and relatively high levels of co-ethnic concentration for the minority identity, we group concentration levels accordingly. Specifically, with respect to the German identity, we employ an indicator variable taking the value of zero – if local co-ethnic concentration in a county is low – and one – if local co-ethnic concentration is medium or high. Similarly, for the minority identity we construct an indicator, which is zero for low or medium levels of local co-ethnic concentration and one for very high levels. The control variables are the same as in our main analysis. We also include again country of origin and county fixed effects, and standard errors reflect within-county and ethnicity category clustering.

The first and third columns in Table 4 contain the OLS regression results for the specification employing the focused binary measures of ethnic concentration; they confirm the findings in columns (2) and (4) in Table 3. Overall, the IV estimates show a broadly similar pattern to the OLS results. The estimates are of analogous magnitude to the OLS estimates. With respect to the German identity (columns (1) and (2)), the IV estimate is a bit smaller in size (-0.377 instead of -0.455), and more imprecisely estimated. The F-statistic of 6.740 in the first stage, indicates a weak instrument, but the parameter estimate (0.179) is highly significant at the 1% level. With respect to the minority identity, (columns (3) and (4)) the IV estimate is practically identical to the OLS estimate (0.221 instead of 0.214), and also statistically significant. The first stage provides again a highly significant parameter estimate (0.348) and a very high F-statistic of 68.415 indicating a strong instrument. The larger standard errors of the IV estimates in comparison to the corresponding OLS results that we observe in Table 4 are in line with general experiences with 2SLS and alternative IV procedures (Wooldridge, 2010, p. 102). With our decent sample size (1, 881) and the very robust and plausible IV parameter estimates, “the weight of statistical evidence should not be primarily assessed on the basis of statistical significance.” (Abadie, 2020, p. 206, and related literature cited there). These findings do not indicate an endogeneity problem and are in line with the other insights reported above.

4.4 Discussion

Our results show that living in an area in which other co-ethnics tend to cluster reduces the likelihood of an immigrant's self-identification with the host country (majority identity) and increases self-identification with the country of origin, in a statistically significant non-linear way. In contrast, residing in a low co-ethnic concentration county increases the likelihood of a majority identity and decreases the country-of-origin identity. Our results are based on an intensive empirical investigation in which core specifications control for a set of demographics as well as pre- and post-migration characteristics. They include both county level and country-of-origin fixed effects and the reported standard errors are cluster-robust at both the country-of-origin and county level. Using the 1970 Full German Census data on co-ethnic spatial dispersion for the first time, we performed a number of robustness checks with IV estimation. Our results confirm stable signs and sizes of the coefficients, albeit with some variations in precision.

Under *cultural conformity* (Bisin et al., 2011, 2016), a high degree of ethnic clustering is likely to strengthen in-group loyalty and immigrants are hence more motivated to retain their respective minority identity and neglect investing in the majority identity. On the other hand, the more sheltered environment of the cluster may reduce the costs of economic and cultural assimilation to the host society, when migrants adjust in groups and not as individuals (Hatton and Leigh, 2011; Aydemir, 2012). Hence ethnic clusters may affect the majority identity formation positively. However, *group assimilation* may not automatically generate rising majority identity in combination with a decline in minority identity, if one allows for multi-ethnic identities. Both the majority and the minority identity may remain strong and co-exist in an immigrant (Constant et al., 2009). Under *cultural distinction*, residing in a low co-ethnic concentration area induces a psychological cost due to the exposure to cultural differences; minority identity helps reduce this cost. Immigrants living in areas with relative high clusters of co-ethnics do not need to preserve their level of minority identity, which might then ease the process of majority identity adaption.

Our empirical results in the historical context of West German guestworker immigration suggest that non-linear ethnic clustering affects the majority identity of immigrants negatively and their minority identity positively, in line with the concept of *cultural conformity*. This rejects both *cultural distinction* (affecting majority identity positively and minority identity negatively) and *group assimilation* (affecting both majority identity and minority identity positively).

5. Conclusions

While ethnic clustering is critically debated in immigrant-receiving societies due to unclear economic and social outcomes, previous research has not studied the effects of ethnic concentration on ethnic identity formation sufficiently well. This crucial omission is not only pertinent to countries with international migration, but it is also pertinent to countries or united countries with internal migration as there are always linguistic, culinary, cultural, and behavioral aspects that characterize certain localities. Internal migration in the U.S. and in China are such examples. At issue is the response of immigrants to the process of technological development and institutional changes which is related to capacity building and living sustainably. How immigrants of different ethnicities and cultures feel about the new host country and the old home country can have important ramifications on their economic, social, and political behavior and integration. Economic development cannot be sustainable if some groups are left behind. We contribute to closing this gap by providing a rigorous analysis using a unique combination of survey data on identity formation with full census data disaggregated at a geographical level.

We study the German paradigm, because Germany is a highly developed country with a long history in migration. West Germany started recruiting labor immigrants in the 1950s and has consistently had a large percentage of immigrants over the last forty years. Empirically, our analysis concerns the historical context of guestworker immigration to West Germany before the unification with East Germany and is based on survey data from the German Socio-Economic Panel for 1987 and the full censuses of 1970 and 1987. Using for the first time data from the unexploited 1987 full German census at a disaggregated geographical level helps us address inherent endogeneity issues. These data allow an accurate calculation of the local ethnic composition at a low, geographically disaggregated level with respect to each ethnic group under study. These groups are Turks, ex-Yugoslavs, Italians, and Greeks.

In addition, we are able to circumvent the issue of immigrant self-selection into certain locations by exploiting the quasi-experimental setting of the historical exogenous geographical distribution of immigrants during the German guestworker scheme of the 1960s and early 1970s. We argue that since migration was purely demand-driven and foreign workers were allocated to specific firms in specific locations before arriving to Germany, the initial geographic placement of immigrants was exogenous to unobserved individual characteristics, such as ethnic identity. Immigration statistics in Germany show

that this locational distribution has been preserved through the years, despite the ban on labor recruitment in 1973 and the subsequent increase in family reunification. Our study thus explores quasi-experimental evidence through the 1970 full German census.

Our analysis shows that living in an area where fellow co-ethnic immigrants tend to cluster reduces the likelihood of self-identification with the host country society. Residential ethnic clustering strengthens immigrants' identification with their respective country of origin. Interestingly, the effects are nonlinear meaning that they become significant only at relatively high levels of co-ethnic concentration for the minority identity and at very low levels of local concentration for the majority identity. Our findings are robust to the use of an instrumental variable approach. These findings can be understood in the context of the *cultural conformity* literature (e.g. Akerlof and Kranton, 2000, 2010), where a high degree of ethnic clustering is predicted to improve in-group loyalties. Therefore, immigrants feel more obliged to keep their minority identity.

Given the large migration experiences Chinese people have historically as well as globally and within China, and the ongoing substantial demographic changes the country observes, knowledge about the formation of ethnic identities and their economic, social and political consequences are of significant importance. The study of diasporas and the benefits and challenges of ethnic clustering can reveal valuable information for policies in host as for sending, in particular developing countries. The consequences of intensifying global ethnic networks for growth, trade, innovations, social, political and international relations are still insufficiently researched.

References

- Abadie, A., 2020. Statistical Nonsignificance in Empirical Economics. *The American Economic Review Insights* 2(2), 193-208.
- Akerlof, G.A., Kranton, R.E., 2000. Economics and identity. *Quarterly Journal of Economics* 115 (3), 715–753.
- Akerlof, G.A., Kranton, R.E., 2010. *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being*. Princeton University Press, Princeton.
- Akay, A., Constant, A.F., Giulietti, C., 2014. The impact of immigration on the well-being of natives. *Journal of Economic Behavior and Organization* 103, 72-92.
- Akay, A., Constant, A.F., Giulietti, C., 2017. Ethnic diversity and well-being. *Journal of Population Economics* 30, 265–306. <https://doi.org/10.1007/s00148-016-0618-8>
- Altschul, I., Oyserman, D., Bybee, D., 2006. Racial-ethnic identity in mid-adolescence: Content and change as predictors of academic achievement. *Child Development* 77 (5), 1155–1169.
- Åslund, O., Edin, P.-A., Fredriksson, P., Grönqvist, P. (2011). Peers, neighborhoods, and immigrant student achievement: Evidence from a placement policy. *American Economic Journal: Applied Economics* 3 (2), 67–95.
- Averett, S.L., Argys, L.M., Kohn, J.C., 2014. Friends with health benefits: Does individual-level social capital improve health? *Eastern Economic Journal*. 40, 181–201.
- Aydemir, A., 2012. Ethnic enclaves and human capital investments: Evidence from a natural experiment, mimeo.
- Battu, H., Zenou, Y., 2010. Oppositional identities and employment for ethnic minorities: Evidence from England. *The Economic Journal* 120 (542), 52–71.
- Bell, B., Machin, S., 2012. Immigrant enclaves and crime. *Journal of Regional Science*. 53 (1), 118–141.
- Bertrand, M., Luttmer, E.F.P., Mullainathan, S., 2000. Network effects and welfare cultures. *Quarterly Journal of Economics* 115 (3), 1019-1055.
- Bisin, A., Patacchini, E., Verdier, T., Zenou, Y., 2016. Bend it like Beckham: Ethnic identity and integration. *European Economic Review* 90, 146–164.
- Bisin, A., Patacchini, E., Verdier, T., Zenou, Y., 2011. Formation and persistence of oppositional identities. *European Economic Review* 55 (8), 1046–1071.
- Bisin, A., Verdier, T., 2023. Advances in the Economic Theory of Cultural Transmission. *Annual Review of Economics* 15 (1), <https://doi.org/10.1146/annurev-economics-090622-100258>.
- Borjas, G.J., 1998. To ghetto or not to ghetto: Ethnicity and residential segregation. *Journal of Urban Economics* 44 (2), 228-253.
- Buch, T., Meister, M., Niebuhr, A., 2021. Ethnic diversity and segregation in German cities. *Cities* 115 (August 2021), 103221, <https://doi.org/10.1016/j.cities.2021.103221>
- Chavous, T.M., Bernat, D., Schmeelk-Cone, K., Caldwell, C., Kohn-Wood, L.P., Zimmerman, M.A., 2003. Racial identity and academic attainment among African American adolescents. *Child Development* 74 (4), 1076–1091.
- Cai, S., Zimmermann, K.F., 2023. Social Identity and Labor Market Outcomes of Internal Migrant Workers. *Global Labor Organization (GLO)*, mimeo.
- Carillo, M.R., Lombardo, V., Venittelli, T. (2023). Social identity and labor market outcomes of immigrants. *Journal of Population Economics* 36 (1), 69–113. <https://doi.org/10.1007/s00148-022->

- Chakraborty, T., Schüller, S., 2022. Ethnic enclaves and immigrant economic integration. *IZA World of Labor* 2022: 287, <https://doi.org/10.15185/izawol.287.v2>.
- Chakraborty, T., S. Schüller, S., Zimmermann, K.F., 2019. Beyond the average: Ethnic capital heterogeneity and intergenerational transmission of education. *Journal of Economic Behavior & Organization* 163 (2019), 551–569.
- Constant, A.F., 1998. The earnings of male and female guestworkers and their assimilation into the German labor market: A panel study 1984 – 1993. Ph.D. Dissertation, Vanderbilt University.
- Constant, A.F., Gataullina, L., Zimmermann, K.F., 2009. Ethnosing immigrants. *Journal of Economic Behavior and Organization* 69 (3), 274-287.
- Constant, A.F., Roberts, R., Zimmermann, K.F., 2009. Ethnic identity and immigrant homeownership. *Urban Studies* 46 (9), 1879–1898.
- Constant, A.F., Zimmermann, K.F., 2009. Work and money: Payoffs by ethnic identity and gender, in: Constant, A.F., Tatsiramos, K., Zimmermann, K.F. (Eds.), *Ethnicity and Labor Market Outcomes*, vol. 29, in: *Research in Labor Economics*, Emerald Group Publishing Limited, Bingley, pp.3–30.
- Constant, A.F., Zimmermann, K.F., 2013. Migration and ethnicity: An introduction, in: Constant, A.F., Zimmermann, K.F. (Eds.), *International Handbook of the Economics of Migration*, Edward Elgar Publishing, Cheltenham, pp.13 – 35.
- Constant, A.F. and Zimmermann, K.F., 2016. Diaspora economics: New perspectives. *International Journal of Manpower* 37 (7), 1-28.
- Corvalan, A., Vargas, M., 2015. Segregation and conflict: an empirical analysis. *Journal of Development Economics* 116, 212–222. <https://doi.org/10.1016/j.jdeveco.2015.05.002>
- Crul, M., 2016. Super-diversity vs. assimilation: how complex diversity in majority–minority cities challenges the assumptions of assimilation. *Journal of Ethnic and Migration Studies*, 42 (1), 54-68, DOI: 10.1080/1369183X.2015.1061425
- Cutler, D.M., Glaeser, E.L., 1997. Are ghettos good or bad? *Quarterly Journal of Economics* 112 (3), 827–872.
- Damm, A.P., 2009. Ethnic enclaves and immigrant labor market outcomes: Quasi-experimental evidence. *Journal of Labor Economics* 27 (2), 281–314.
- Danzer, A.M., Feuerbaum, C., Piopiunik, M., Woessmann, L. (2022). Growing up in ethnic enclaves: language proficiency and educational attainment of immigrant children. *Journal of Population Economics* 35 (4), 1297–1344. <https://doi.org/10.1007/s00148-022-00889-y>
- Danzer, A.M., Yaman, F., 2013. Do ethnic enclaves impede immigrants’ integration? Evidence from a quasi-experimental social-interaction approach. *Review of International Economics* 21 (2), 311–325.
- Danzer, A.M., Yaman, F., 2016. Ethnic concentration and language fluency of immigrants: Evidence from the guest-worker placement in Germany. *Journal of Economic Behavior and Organization* 131, 151–165.
- Edin, P.-A., Fredriksson, P., Aslund, O., 2003. Ethnic enclaves and the economic success of immigrants - Evidence from a natural experiment. *Quarterly Journal of Economics* 118 (1), 329–357.
- Ehrkamp, P., 2005. Placing identities: Transnational practices and local attachments of Turkish immigrants in Germany. *Journal of Ethnic and Migration Studies* 31 (2), 345–364.
- Fuligni, A.J., 2011. Social identity, motivation, and well being among adolescents from Asian and Latin American backgrounds. In G. Carlo, N. J. Crockett, & Carranza, M. (Eds.) *Health*

Disparities in Youth and Families: Research and Applications. Nebraska Symposium on Motivation, Vol. 57, 97-120. New York, NY: Springer Science.

Glitz, A., 2014. Ethnic segregation in Germany. *Labour Economics* 29, 28–40.

Goebel, J.; Grabka, M. M.; Liebig, S.; Kroh, M.; Richter, D.; Schröder, C.; Schupp, J. (2019): The German Socio-Economic Panel Study (SOEP), *Jahrbücher für Nationalökonomie und Statistik / Journal of Economics and Statistics* 239 (2): 345–360.

Han, E.; Paik, C. (2017) “Ethnic Integration and Development in China.” *World Development*, 93, 31–42, <http://dx.doi.org/10.1016/j.worlddev.2016.12.010>

Hanke, T. 2023. Parallel-Gesellschaften durch Zuwanderer? Der Heiratsmarkt regelt das (Parallel societies through immigrants? The marriage market regulates that). *Handelsblatt*. <https://www.handelsblatt.com/meinung/kolumnen/globale-trends-parallel-gesellschaften-durch-zuwanderer-der-heiratsmarkt-regelt-das/29064160.html>

Hatton, T.J., Leigh, A., 2011. Immigrants assimilate as communities, not just as individuals. *Journal of Population Economics* 24 (2), 389–419.

Hernandez, D.J., Macartney, S. Blanchard, V.L., 2010. Children of immigrants: Family and socioeconomic indicators for affluent countries. *Child Indicators Research* 3 (4), 413–437.

Heider, B., Stoms, P., Koch, J., Siedentop, S. (2020). Where do immigrants move in Germany? The role of international migration in regional disparities in population development, *Population, Space and Place*, 26 (8) <https://doi.org/10.1002/psp.2363>

Koczan, Z., 2012. So, what are you? And does this matter? Second generation identities: Formation and effects. Draft conference paper, accessed in September 2014 from: <http://www.inter-disciplinary.net/wp-content/uploads/2012/02/koczanpaper.pdf>

Layard, R., 2005. *Happiness: Lessons from a new science*. London: Penguin Books.

Mägi, K., van Ham, M., Leetmaa, K., Tammaru, T. (2020). The neighbourhood context and changes in self-reported ethnic identity, *Journal of Ethnic and Migration Studies*, 46:1, 277-296, DOI: 10.1080/1369183X.2018.1547634

Mahlzahn, C. C., 2023. Wie versteinerte Parallelgesellschaften Berlin zerfleddern (How fossilized parallel societies are tearing Berlin apart). *DIE WELT*. <https://www.welt.de/politik/deutschland/-plus243782209/Berlin-Wie-Parallelgesellschaften-die-Hauptstadt-zerfleddern.html>

Massey, D. S., 1985. Ethnic residential segregation: A theoretical and empirical review. *Sociology and Social Research* 69 (3), 315–350.

Matuszeski, J., Schneider, F., 2006, Patterns of ethnic segregation and civil conflict, Mimeo, Harvard University.

Oyserman, D., Yoon, K.-I., 2009. Neighborhood effects on racial-ethnic identity: The undermining role of segregation. *Race and Social Problems* 1, 67–76.

Pendakur, K., Pendakur, R., 2005. Ethnic identity and the labour market, Mimeo, Simon Fraser University.

Piracha, M., Tani, M., Cheng, Z., Wang, B. Z., 2023. Social assimilation and immigrants’ labour market outcomes. *Journal of Population Economics* 36 (1), 37–67. <https://doi.org/10.1007/s00148-021-00883-w>

Schönwälder, K., Söhn, J., 2009. Immigrant settlement structures in Germany: General patterns and urban levels of concentration of major groups. *Urban Studies* 46 (7), 1439–1460.

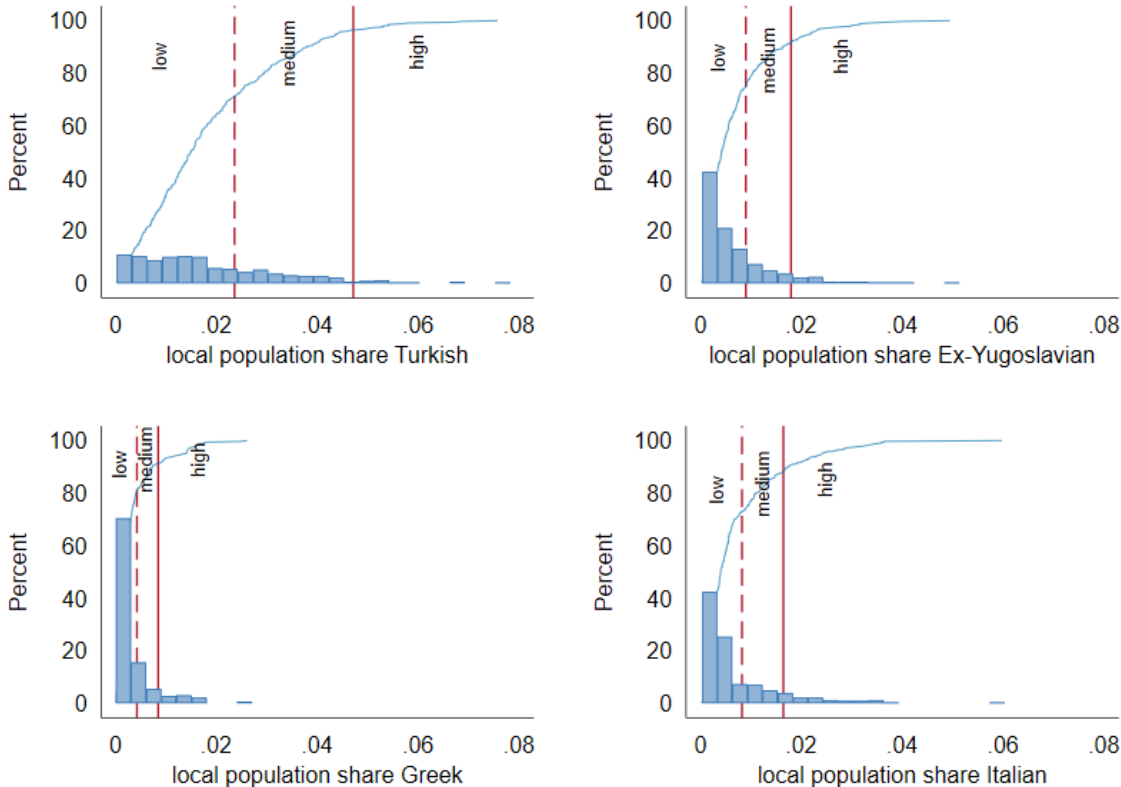
Statistisches Bundesamt, 2010. Bevölkerung und Erwerbstätigkeit: Bevölkerung mit Migrationshintergrund – Ergebnisse des Mikrozensus 2010. Fachserie 1, Reihe 2.2.

U.S. Census Bureau. 2012. U.S. Census Bureau projections show a slower growing, older, more diverse nation a half century from now. Release CB12-243. Retrieved from: <https://>

www.census.gov/newsroom/releases/archives/population/cb12-243.html (last accessed on 29.10.2014).

- Winke, T. (2018). Being close to ethnic peers: the formation of ethnic residential clustering from a household perspective, *Journal of Ethnic and Migration Studies*, 44 (11), 1932–1953. <https://doi.org/10.1080/1369183X.2017.1368373>
- Xu, D., Zhang, Y. (2022). Identifying ethnic occupational segregation. *Journal of Population Economics* 35 (3), 1261–1296 (2022). <https://doi.org/10.1007/s00148-020-00796-0>
- Zimmermann, L., Gataullina, L., Constant, A.F., Zimmermann, K.F., 2008. Human capital and ethnic self-identification of immigrants. *Economics Letters* 98 (3), 235–239.
- Zhang, Y., Xu, D. (2023). Ethnic Occupational Segregation. In: Zimmermann, K.F. (ed) *Handbook of Labor, Human Resources and Population Economics*. Springer, Cham. https://doi.org/10.1007/978-3-319-57365-6_414-1
- Zwysen, W., Demireva, N. (2020). Ethnic and migrant penalties in job quality in the UK: the role of residential concentration and occupational clustering, *Journal of Ethnic and Migration Studies*, 46:1, 200-221, DOI: 10.1080/1369183X.2018.1498777
- Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*, The MIT Press, Cambridge, Massachusetts & London, England.

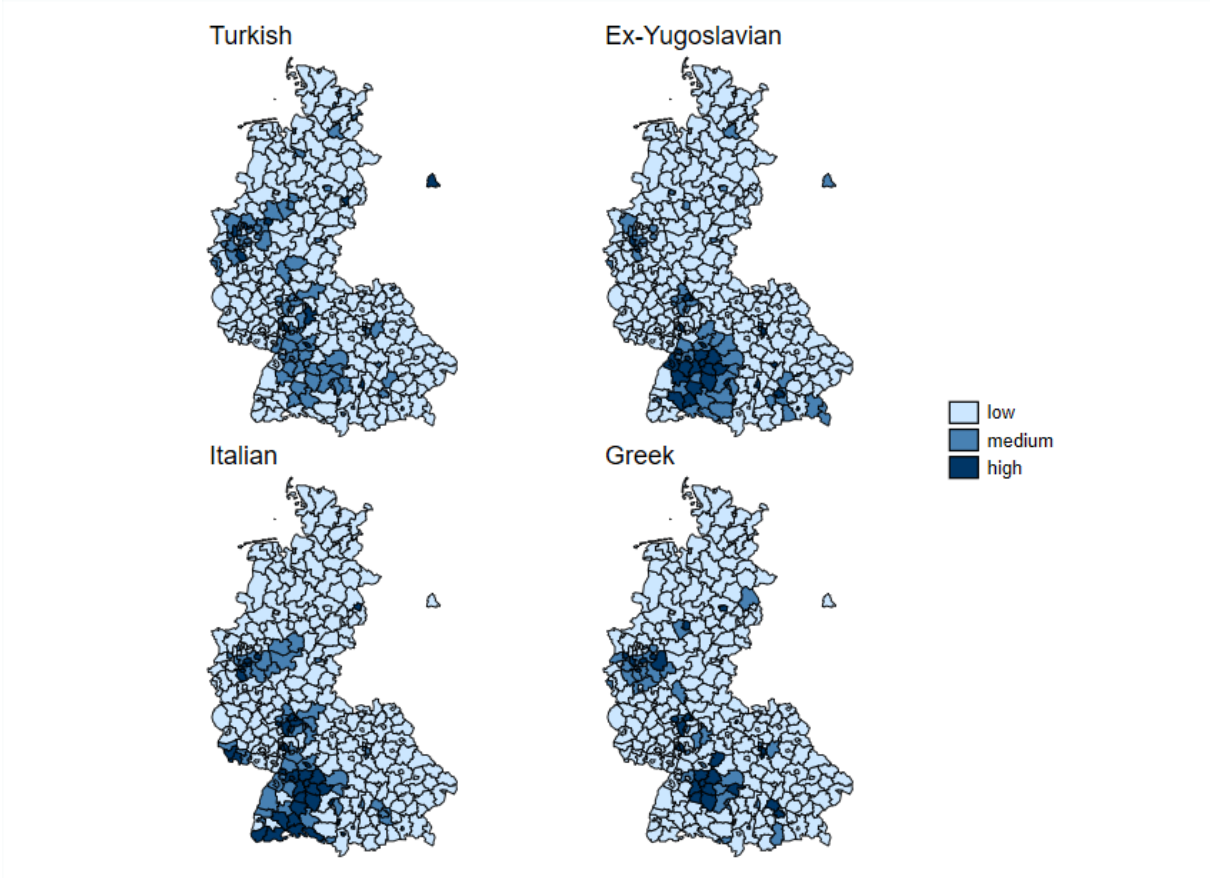
Figure 1
 Cumulative and absolute distribution of county-level ethnic concentration by ethnic group.



Notes: The dashed vertical line indicates the national population share of the respective ethnic group. The solid vertical line represents twice that share.
Source: Research Data Centres of the Federal Statistical Office and the Statistical Offices of the Länder, Full German Census 1987, West Germany, own calculations. County-level information.

Figure 2

Geographic concentration of selected ethnic groups across the former West Germany.



Notes: County-level information. A low (medium, high) co-ethnic concentration is defined as a county with a local co-ethnic fraction which is lower (at least as high, at least twice as high) as the national population share of the respective group. Darker areas denote higher co-ethnic concentration counties.

Source: Research Data Centres of the Federal Statistical Office and the Statistical Offices of the Länder, Full German Census 1987, West Germany, own calculations.

Table 1
Ethnic concentration in Germany

	German Citizenship	Foreign Citizenship					
		Total	Turkey	The Former Yugoslavia	Italy	Greece	Other
<i>A. Full German Census 1987</i>							
Total number of individuals	56,029,672	4,081,959	1,415,425	538,707	493,022	253,433	1,381,372
National population share (%)	93.21	6.79	2.35	0.90	0.82	0.42	2.30
Mean local ethnic group size	170,822 (161,740)	12,445 (22,246)	4,315 (8,763)	1,642 (3,821)	1,503 (2,505)	773 (1,724)	2,764 (5,676)
Mean local population share (%)	94.67 (3.50)	5.33 (3.50)	1.82 (1.39)	0.67 (0.73)	0.70 (0.81)	0.30 (0.41)	1.20 (0.84)
Max local share in one county (%)	99.43	20.40	7.58	4.93	5.97	2.59	5.83
Medium local co-ethnic concentration (%)			49.97	33.37	26.83	19.85	
High local co-ethnic concentration (%)			21.19	40.04	45.91	50.20	
Top-5 counties							
	1		Duisburg	Stuttgart	Wolfsburg	Offenbach/M.	
	2		Gelsenkirchen	Frankfurt/M.	Waldshut	Stuttgart	
	3		Köln	München	Lörrach	Rems-Murr-Kreis	
	4		Herne	Offenbach/M.	Solingen	Ludwigshafen/R.	
	5		Berlin/W.	Calw	Ludwigshafen/R.	Dachau	
<i>B. SOEP Sample 1987</i>							
Mean local population share (%)	92.32 (3.66)	7.68 (3.66)	2.68 (1.40)	1.07 (0.89)	1.11 (1.01)	0.51 (0.52)	
Medium local co-ethnic concentration (%)			54.73	38.11	20.35	14.95	
High local co-ethnic concentration (%)			20.88	34.95	58.04	57.48	

Source: Research Data Centres of the Federal Statistical Office and the Statistical Offices of the Länder, Full German Census 1987, West Germany, own calculations.

Note: Local level is county level (total 328 counties). Standard deviations in parentheses. Ethnic concentration is the size of the ethnic group relative to the population in each county. A low (medium, high) co-ethnic concentration is defined as a county with a local co-ethnic fraction which is lower (at least as high, at least twice as high) as the national population share of the respective group.

Table 2
Descriptive statistics of main variables by co-ethnic concentration.

	Co-Ethnic Concentration					
	Low		Medium		High	
	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
German Identity (1-5)	1.936	(1.045)	1.921	(1.050)	2.047	(1.075)
Minority Identity (1-5)	4.406	(0.881)	4.247*	(1.004)	4.232*	(0.982)
Female	0.459	(0.499)	0.454	(0.498)	0.452	(0.498)
Age at Migration	22.983	(10.567)	22.295	(10.315)	22.295	(10.621)
Turkey	0.402	(0.491)	0.599*	(0.491)	0.227*	(0.419)
The Former Yugoslavia	0.237	(0.426)	0.223	(0.416)	0.203	(0.403)
Italy	0.184	(0.388)	0.115*	(0.319)	0.325*	(0.469)
Greece	0.177	(0.382)	0.064*	(0.245)	0.244*	(0.430)
<i>Pre-migration education</i>						
no schooling in home country	0.205	(0.404)	0.177	(0.382)	0.175	(0.380)
college in home country	0.013	(0.113)	0.014	(0.118)	0.018	(0.134)
vocational training in home country	0.265	(0.442)	0.247	(0.431)	0.220	(0.415)
complete schooling in home country	0.306	(0.461)	0.309	(0.463)	0.277	(0.448)
incomplete schooling in h. country	0.212	(0.409)	0.252	(0.435)	0.309*	(0.463)
<i>Post-migration education</i>						
some schooling in Germany	0.652	(0.477)	0.662	(0.473)	0.612	(0.488)
no degree in Germany	0.340	(0.474)	0.336	(0.473)	0.378	(0.485)
higher degree in Germany	0.009	(0.092)	0.001	(0.038)	0.010	(0.099)
Years since Migration	15.506	(5.653)	14.833*	(5.459)	16.740*	(5.906)
Employment spell in Germany	0.876	(0.330)	0.840	(0.367)	0.887	(0.317)
Local unemployment rate	9.228	(3.380)	9.283	(3.942)	8.514*	(3.776)
Local native population density/1000	0.501	(0.708)	1.158*	(0.911)	1.463*	(1.054)
Ln(local immigrant group size)	8.999	(0.996)	9.964*	(1.000)	10.358*	(1.067)
Nr. of Observations	468		705		708	

Source: SOEP 1987; Research Data Centres of the Federal Statistical Office and the Statistical Offices of the Länder, Full German Census 1987, West Germany, own calculations.

Note: Standard deviations in parentheses. Local level is county level. Ethnic concentration is the size of the ethnic group relative to the population in each county. A low (medium, high) co-ethnic concentration is defined as a county with a local co-ethnic fraction which is lower (at least as high, at least twice as high) as the national population share of the respective group. * Statistically different from low co-ethnic concentration mean at the 5 percent confidence level.

Table 3
Ethnic identity and co-ethnic concentration.

	German Identity		Minority Identity	
	(1)	(2)	(3)	(4)
Ln(local ethnic group size)	-0.041 (0.085)		0.118 (0.080)	
<i>Low local ethnic concentration (ref.)</i>				
Medium local ethnic concentration		-0.443*** (0.137)		0.137 (0.103)
High local ethnic concentration		-0.476*** (0.148)		0.334*** (0.109)
Female	-0.063 (0.043)	-0.063 (0.043)	0.018 (0.040)	0.019 (0.040)
<i>No schooling in home country (ref.)</i>				
College in h.c.	0.380** (0.186)	0.386** (0.184)	-0.343 (0.223)	-0.338 (0.220)
Vocational training in h.c.	0.149* (0.087)	0.149* (0.086)	-0.092 (0.090)	-0.095 (0.090)
Complete schooling in h.c.	-0.035 (0.087)	-0.037 (0.086)	0.085 (0.091)	0.083 (0.091)
Incomplete schooling in h.c.	0.104 (0.072)	0.106 (0.072)	-0.039 (0.073)	-0.040 (0.073)
<i>Some schooling in Germany (ref.)</i>				
No degree in G.	-0.158** (0.077)	-0.158** (0.078)	0.154* (0.087)	0.152* (0.087)
Higher degree in G.	0.667*** (0.158)	0.614*** (0.153)	-0.574*** (0.196)	-0.573*** (0.194)
Employment spell in G.	0.131 (0.087)	0.125 (0.086)	-0.110 (0.076)	-0.110 (0.076)
Years since migration	0.013*** (0.005)	0.013*** (0.005)	-0.012** (0.005)	-0.012** (0.005)
Age at migration	-0.041*** (0.010)	-0.042*** (0.010)	0.038*** (0.009)	0.038*** (0.009)
Age at migration squared/1000	0.391** (0.167)	0.403** (0.167)	-0.380** (0.166)	-0.381** (0.166)
<i>Turkish (ref.)</i>				
Ex-Yugoslavian	0.271** (0.111)	0.288*** (0.082)	-0.075 (0.115)	-0.188** (0.088)
Italian	0.119 (0.114)	0.141 (0.090)	0.045 (0.119)	-0.109 (0.107)
Greek	-0.298** (0.146)	-0.287*** (0.104)	0.481*** (0.145)	0.252** (0.107)
Constant	0.693 (0.811)	0.681*** (0.173)	-1.470** (0.745)	-0.541*** (0.146)
<i>County FE</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
<i>N</i>	1,881	1,881	1,881	1,881
<i>adj. R-sq</i>	0.250	0.255	0.198	0.200
<i>AIC</i>	4690.913	4681.430	4817.212	4813.820
<i>BIC</i>	4779.546	4775.602	4905.845	4907.993

Source: SOEP 1987; Research Data Centres of the Federal Statistical Office and the Statistical Offices of the Länder, Full German Census 1987, West Germany, own calculations.

Notes: Standard errors, in parentheses, are adjusted to reflect within-county/ethnicity clustering. Local level is county level. The dependent variables are measured on a five-point scale from “no” (1) to “full” (5) identification with the host (home) country and here included as a standardized quasi-metric measure. A low (medium, high) ethnic concentration is defined as a county with a local co-ethnic fraction which is lower (at least as high, at least twice as high) as the national population share of the respective ethnic group. * $p < .10$; ** $p < .05$; *** $p < .01$

Table 4
Instrumental variable estimation.

	German Identity		Minority Identity	
	(1)	(2)	(3)	(4)
	OLS	IV	OLS	IV
Medium/high local ethnic concentration	-0.455***	-0.377	0.214**	0.221*
	(0.131)	(0.494)	(0.092)	(0.124)
1970 medium/high local ethnic concentration				
<i>Controls</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
<i>Country of origin FE</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
<i>County FE</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
<i>First stage:</i>				
1970 medium/high local ethnic concentration		0.179***		0.348***
		(0.069)		(0.042)
Partial R-sq.		0.890		0.880
F-statistic		6.740		68.415
<i>N</i>	1,881	1,881	1,881	1,881

Source: SOEP 1987; Research Data Centres of the Federal Statistical Office and the Statistical Offices of the Länder, German census 1987 and 1970, West Germany, own calculations.

Notes: Minority groups covered are Turks, former Yugoslavians, Italians and Greeks. Standard errors in parentheses with within-county and ethnicity clustering. Local level is the county level. The dependent variables are measured on a five-point scale from “(1) to “full” (5) identification with the host (home) country and here included as a standardized quasi-metric measure. A low (medium, high) ethnic concentration is defined as a county with a local co-ethnic fraction which is lower (at least as high, at least twice as high) as the national population share of the respective group. In IV models, local ethnic concentration measure is instrumented using the respective co-ethnic concentration in 1970 employing the linear binary probability model in the first stage.

* $p < .10$; ** $p < .05$; *** $p < .01$