Does Unprecedented Mass Immigration Fuel Ethnic Discrimination? A Two-Wave Field Experiment in the German Housing Market

Katrin Auspurg, Renate Lorenz, Andreas Schneck

Ludwig Maximilian Universität Munich

Abstract: Literature suggests that sudden mass immigration can fuel xenophobic attitudes. However, there is a lack of reliable evidence on hostile actions, such as discrimination. In this study, we leverage the unexpected mass immigration of refugees to Germany in 2015 in combination with a two-wave field experiment to study the effect of immigration on ethnic discrimination. In 2015/2016, political and social tensions in the Middle East and North Africa led to a historic mass migration to European countries. We carried out a large-scale field experiment on ethnic housing market discrimination in Germany (paired e-mail correspondence test with ~5,000 e-mail applications to rental housing units in each wave) shortly before this "European refugee crisis"(1st wave). We repeated this experiment at the peak of the crisis (2nd wave of our experiment). By taking advantage of the unexpected refugee immigration between the two waves of our experiment and the quasi-random allocation of refugees across regions for causal identification, we find no credible evidence that the large influx of refugees changed the extent of ethnic discrimination of Turks in the rental housing market. This result holds regardless of the extent to which regions within Germany were already accustomed to immigration before the refugee crisis.

Keywords: immigration; ethnic discrimination; housing market; field experiment; segregation; asylum

IN recent years, the number of refugees has continuously increased, with the share of displaced persons accounting for more than 1 percent of the world's population (1 in 88 people in 2021; UNHCR 2022). Due to violence and conflicts (such as Russia's war of aggression against Ukraine), human rights violations, political instability, demographic change, or severe impacts of climate change, refugee migration is expected to remain a global mega-trend in the future (Kaczan and Orgill-Meyer 2020, Koubi 2019).

In this study, we ask whether unexpected mass immigration of refugees affects the discrimination of immigrants already living in a host country. A common assumption is that the majority population perceives an influx of immigrants with a different ethnic or religious background as culturally threatening (see, e.g., Brady and Finnigan 2014, Hainmueller and Hopkins 2014, Newman 2013). We expect that resulting negative feelings against refugees spill over to Turkish immigrants already living in Germany. There is evidence that boundaries for defining threatening groups in Europe are drawn primarily along religion, not ethnicity (for an overview: Gereke et al. 2022). At the same time, it is to be expected that both immigrating refugees and Turks are predominantly perceived as Muslims. An increase in the relative size of immigrants (out-group) may also increase competition (real or

perceived) between immigrants (out-group) and natives (in-group) for social and economic resources. A sense of (competitive) threat is likely to increase hostility and motivation to discriminate (e.g., Blalock 1967, Hopkins 2010, Semyonov et al. 2004). All of these mechanisms can jeopardize the integration success of newly arriving refugees, but they can also lead to setbacks for immigrants already living in a country. Although research often shows that immigration can increase anti-migration sentiments and support for anti-immigration policies (e.g., Hopkins 2010, Semyonov et al. 2004), there is little research on the impact of immigration on discrimination.

In our study, we therefore pursue the following questions: (i) Did the strong immigration during the European refugee crisis increase discrimination in the German rental housing market against Turkish immigrants? (ii) Are there heterogeneous effects depending on regions' prior exposure to immigrant populations?

Our major contribution to the literature is to exploit the unexpected immigration of refugees from predominantly African and Middle Eastern countries in late summer and fall 2015, known as the European "refugee" or "migration crisis," to identify the effect of immigration on discrimination. Discrimination was measured with a paired field-experimental design that ensures high internal validity (Baldassarri and Abascal 2017, Elwert, Keller and Kotsadam 2023). For the first research objective, identifying the average treatment effect of immigration on discrimination, we use the highly salient and unanticipated event of the refugee crisis as a kind of natural experiment. Our repeated field experiment consists of two waves, with the 1st wave conducted in May before the unexpected immigration (control group), and the 2nd wave conducted in December after the unexpected immigration unfolded (treatment group). Repeating the field experiment with the same design and in the same regional housing markets allows for balancing out most covariates. In addition, the mass immigration was unforeseen, thus preventing anticipation effects: Only the treatment group of housing suppliers tested in the 2nd wave was exposed to the refugee crisis. With certain further assumptions about exogeneity (Muñoz, Falcó-Gimeno and Hernández 2020, evaluated later), this design allows identifying the causal effect of refugee immigration on ethnic housing discrimination of Turks.

For our second research goal, identifying treatment effect heterogeneity by regions differentially accustomed to immigration, we also make use of the specific setting of the refugee crisis. Germany's legal regulations ensured a quasi-random allocation of refugees to geographic regions (depending on population size and tax revenues) and at least prevented refugees from self-selecting into regions with lower discrimination or higher proportions of foreigners. This exogeneity (also discussed in more detail later) allows us to identify treatment heterogeneity without confounding or endogeneity bias that occurred in previous (field-experimental) studies (for general threats to identification in experimental research: Montgomery, Nyhan and Torres 2018, Vanderweele 2015).

To preview our results: The discrimination measured with our field experiments was remarkably immune to the influence of the refugee crisis: The level of discrimination observed at the peak of the refugee immigration (2nd wave) did not differ substantially from what we observed shortly before the refugee crisis unfolded (1st wave). Throughout, e-mail applications from Turkish applicants had

a 10 percentage points lower chance of receiving a response from housing suppliers than the same e-mail applications from German applicants. This result applies regardless of the extent to which regions were already accustomed to immigration before the refugee crisis.

Background

Theories on Immigration, Group Threat and Discrimination

Theories on threats and intergroup conflicts predict that prejudice and hostile attitudes against immigrants—which are both seen as predictors for discrimination (Becker 1957)—increase with the (perceived) number of immigrants relative to the autochthonous population in a geographic region (Blalock 1967, for evidence on the U.S.: Newman 2013, Ha 2010, Hopkins 2010, for evidence on Europe: Markaki and Longhi 2013, Semyonov et al. 2004). Supposed reasons are that immigrants are perceived as culturally different (e.g., because they have different values or a different religion) and that an increase in their relative size is perceived as a cultural or economic threat to the majority culture (e.g., because an increased number of immigrants would have more political and cultural influence or could exploit the welfare system to a greater extent; Brady and Finnigan 2014). Rising feelings of threat could affect real estate agents and private suppliers in the housing market as part of the majority population in Germany.[1] The influx of refugees may also threaten rental income and property values, as a decline in rents has already been recorded in the neighborhood of refugee accommodations (for evidence on the European refugee crisis: Hennig 2021, Kürschner Rauck 2020). Another possible reason for immigration fueling anti-immigration sentiment is that a growing immigrant population intensifies competition between natives and immigrants over scarce resources, such as labor market opportunities or childcare facilities. Although homeowners generally belong to a wealthier population segment, their social networks may not be fully insulated from perceived competition. In addition, because real estate agents do not require a formal qualification in Germany, agents with low education may feel threatened by increased competition for jobs, housing or social benefits. Natives, including housing suppliers, may also fear greater ethnic mixing and resulting declines in the perceived quality of their children's schools (Betts and Fairlie 2003).

Regardless of what causes the feelings of threat: We expect that exposure to immigration spurs exclusionary attitudes on the part of the ethnic majority toward their ethnic others (Elwert, Keller and Kotsadam 2023). Indeed, there is a large body of literature based on theories of ethnic threat and competition (Blalock 1967, Blumer 1958) that shows that increases in immigrant populations enhance anti-immigrant sentiments. In particular, sudden mass immigration is likely to lead to threats and competition, as it limits the time for the population to adapt to cultural change (e.g. by positive contact experiences with immigrants) and for markets to absorb the increasing demand (Cea D'Ancona 2018, Olzak 1994). This is consistent with the two necessary conditions identified by Hopkins (2010) for turning native attitudes against immigrants: Mass immigration noticeably shifts the sociodemographic

composition at the local level, and at the same time, there is a salient national rhetoric about the threat of immigration. Presumably, a large influx of immigrants is particularly culturally threatening to citizens living in environments with little previous exposure to immigrants (Newman 2013).

The setting of the European refugee crisis fits these scope conditions very well and thus might be one of the most likely cases of threat and conflict theories to apply. The influx of refugees was unexpected and exceptionally strong, even for a country like Germany, which had been accustomed to immigration (see statistics in later sections). In Eastern Germany, refugees also settled in regions that had previously hardly been exposed to immigration: The share of foreigners (population without German citizenship) at the time of the refugee crisis was below 1 percent in several Eastern German municipalities. The immigrants were hosted mainly in special homes, so-called initial reception facilities after their arrival. However, immigration also intensified competition in the private housing market, as shelters were supposed to be temporary and refugees were supposed to move into (subsi-dized) private housing when they found such accommodation (Hennig 2021). At the same time, the refugee crisis became the primary topic of public, political, and media debates in Germany and Europe (Wagner et al. 2020). After initial reports on Germany's exceptional "welcome culture," the media increasingly focused on integration challenges, protests against immigration, and hostile attacks on refugee shelters (Czymara and Schmidt-Catran 2017, Jäckle and König 2018, Wagner et al. 2020). Thus, we expect theories on group conflicts (and not opposing contact theories[2]) to apply: The refugee immigration has most likely intensified natives' out-group rejection.

What is more difficult to predict is which immigrant groups will be affected: only newly arrived immigrants or also those who have lived in the host country for some time? To date, there are conflicting ideas about how the arrival of new immigrants can change perceptions and behavior toward immigrants already living in the host country (Fouka and Tabellini 2021). Former immigrants may benefit from the arrival of new immigrants by being re-classified as a group less culturally distant from the native population, for example, due to comparatively better language proficiency. Such positive reframing of the perception of one out-group due to the appearance of a new out-group was observed in the U.S., where the arrival of Mexican immigrants lowered prejudice against Black Americans (Fouka and Tabellini 2021). In the case of the European refugee crisis, however, we expect that the negative out-group rejection caused by immigrating African and Middle Eastern refugees will spread to other Muslim minority groups, including the Turkish migrants tested in our experiments. Such negative cultural sociotropism was found, for example, in the aftermath of the terror attacks on 9/11, when anti-Muslim rhetoric sparked a backlash against all immigrant groups (Hopkins 2010). Spillover effects are considered more likely if a minority group is perceived as culturally similar to new immigrants, which may lead the native population to lump this group together with the new immigrants as one socially distant out-group. In Europe, being Muslim or not seems to be an even more important predictor for group boundaries definition, discriminatory attitudes and behavior than race or ethnic origin (Alba 2005, Auspurg, Schneck and Hinz 2019, Di Stasio et al. 2021,

Gereke et al. 2022). Many Europeans and Germans perceive a strong incompatibility between Muslim and Western values (Korteweg and Yurdakul 2009, Zolberg and Woon 1999). Turkish migrants were therefore likely to be perceived as culturally close to refugees who immigrated primarily from Syria, Iraq and Afghanistan, as all these migrants have an origin in a Muslim country (for more detailed arguments: Deole and Huang 2020).

Research on Attitudes

With respect to the European refugee crisis, there is evidence that the large influx of refugees has fostered anti-immigration attitudes and sympathy with far-right-parties opposing immigration (Czymara and Schmidt-Catran 2017, Dinas et al. 2019, Rudolph and Wagner 2022, Steinmayr 2021, for evidence on Germany: Dostal 2017, Czymara 2021, for an exception: Schaub, Gereke and Baldassarri 2020). Some studies have also reported spillover effects on the Muslim population in Europe. For example, exposure to refugees near reception centers made locals in Greece more hostile to Muslims who had lived on these Greek islands for generations (Hangartner et al. 2019).

Studies on attitudes have justified their focus by arguing that attitudes motivate hostile actions (Dinas et al. 2019, Semyonov et al. 2004). However, this argument misses a crucial point: Ethnic discrimination is arguably not just mapping anti-immigrant attitudes into actions. Instead, discrimination is regulated by responses to perceived norms of appropriate behavior and cost-benefit considerations (Böhm et al. 2018, Scacco and Warren 2018). In the housing market, higher search costs result when landlords accept only one ethnic group as possible tenants (Auspurg, Schneck and Thiel 2020). Lab experiments have shown that even low-cost burdens can prevent anti-immigration sentiments from resulting in corresponding hostile actions (Böhm et al. 2018). Furthermore, widely studied statistical and customer discrimination theories (Arrow 1971, Phelps 1972) suggest that, instead of feelings of dislike, seeking stable rental payments in combination with assumptions about prospective tenants' solvency may be the cause of ethnic discrimination. Many natives see housing in or near ethnic enclaves and refugee shelters as a disamenity (Hennig 2021, Liebe et al. 2018). This could even prompt landlords to favor immigrant applicants over native applicants in multiethnic neighborhoods, as steering migrants to other migrants would reduce the time it takes real estate owners and agencies to rent housing units (Farley et al. 1994).

In sum, it can be expected that unforeseen mass immigration of refugees will fuel prejudices and anti-migration sentiments; this may be especially true in regions that are barely accustomed to immigration. Still, it is unclear if this translates to increased ethnic discrimination against immigrants already living in a country.

Assessing Effects of Immigration on Discrimination

A body of literature shows that discrimination on (housing) markets still constrains housing and neighborhood opportunities for ethnic or religious minorities (for meta-analyses: Auspurg, Schneck and Hinz 2019, Flage 2018, Quillian et al. 2017). Many studies have shown that discrimination in the housing market can have

detrimental effects on minority groups' short- and long-term outcomes (e.g. by negatively affecting their education, labor market, and/or health status; for an overview: Krysan and Crowder 2017). Therefore, it is not surprising that a bulk of studies tried to identify conditions under which discrimination most likely occurs, including the ethnic composition of neighborhoods and variations in the size of immigrant populations.

To date, nearly all empirical studies on the immigration-discrimination nexus have studied correlations between the actual size or share of immigrant populations in different regions and the level of discrimination observed there.[3] So far, the results from these studies are inconclusive. Although most studies on the U.S. found the level of discrimination to decrease with the share of immigrants living in a region (for an overview: Krysan and Crowder 2017), the evidence on Europe is mixed. Some studies found no (e.g. Van der Bracht, Coenen and Van de Putte 2015), others a positive (e.g. Baldini and Federici 2011, Auspurg, Hinz and Schmid 2017), and some a negative association (e.g. Carlsson and Eriksson 2014) between the regional size of immigrant populations and discrimination.

This heterogeneous state of research probably results from an empirical research design that does not allow for causal identification. To interpret the statistical effect of the share of immigrants as a causal effect, one must assume that unobserved factors that simultaneously affect immigration and discrimination (such as the tightness of housing and labor markets) are invariant across the compared regional units. Especially when comparing aggregated data, there is a risk of bias due to omitted variables and ecological fallacy. Furthermore, there is the problem of self-selection (Krysan and Crowder 2017), as individuals can often decide whether or not they want to interact with other groups (as neighbors). It is plausible that individuals with a priori xenophobic attitudes prefer real estate ownership more often in regions with fewer members of other ethnic/religious groups than individuals with openness to multiethnic neighborhoods. At the same time, immigrants may self-select into specific regions (e.g., with higher numbers of co-ethnics) to evade discrimination. In case of such self-selection, the causality would be reversed: discrimination causing low immigration, not vice versa.

Figure 1 illustrates these various correlations between immigration and discrimination. Panel (A) shows a causal effect of immigration on discrimination, mediated by increasing cultural threats and competition. In panel (B), the supposed causal effect is only spurious, caused by confounding variables and by self-selection of property owners and immigrants. We contribute to the literature by seeking to identify a causal effect of immigration, as hypothesized in panel (A). This is possible because, unlike previous research, we can draw on exogenous variation in the size of immigrant populations. Even though we cannot infer the mediating variables (we have no information on feelings of threat and competition at the micro level of rental processes), this is an important first step in clarifying the causal structure underlying the frequently observed association between immigration and discrimination.

We are aware of only one study focusing on discrimination in the context of the European refugee crisis. This lab experiment compared regions in Eastern Germany with varying exposure to refugees and concluded that the (strong) influx of refugees
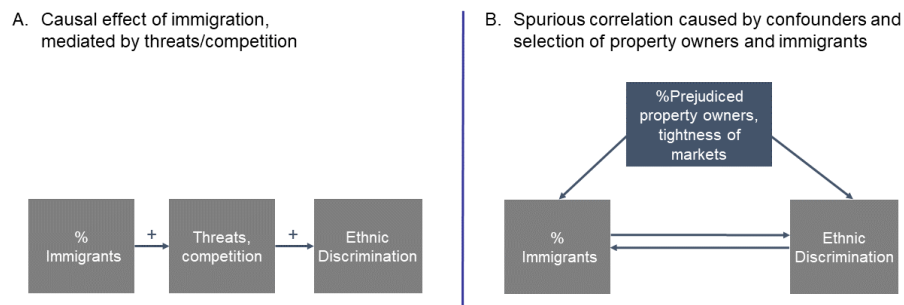
Figure 1: Possible Associations between Immigration and Discrimination.
Note: In panel (A), one can expect throughout positive associations, if threat or conflict theories apply. In panel (B), we do not show the signs of associations, as confounders might have positive or negative effects on immigration and discrimination. Therefore—if not considered—confounders can lead to an over- or underestimation of the effect of immigration on discrimination.

did not alter hostile attitudes, voter behavior, or discrimination (Schaub, Gereke and Baldassarri 2020). However, lab experiments might show limited external validity due to selective participation and limited internal validity due to the artificiality of lab experimental settings (Shadish, Cook and Campbell 2002). In addition, the cross-sectional setting of this study did not allow for observing potential changes in discrimination over time.

Data and Research Design

The Setting of the Refugee Crisis

In 2015, around 1.3 million people applied for asylum in Europe—more than in every single year since the Second World War. Germany alone took in nearly one million refugees, most of whom were Muslim (German Federal Office for Migration and Refugees - BAMF 2016). Although there had been a small increase in refugee numbers since 2011, a sudden and unexpected mass influx was triggered by the decision of German authorities in August 2015 to suspend the "Dublin Agreement"[4] for Syrian refugees (and one month later also for refugees from other countries). Refugees were allowed to file their asylum application in Germany from then on, even if the country of arrival would normally have been responsible for processing the application.

Even though Germany has had a long history of receiving refugees, the subsequent uptick in immigration in the second half of 2015 was exceptional.[5] This can be seen in Figure 2, which plots the monthly registration of asylum seekers together with the timing of the two waves of our repeated field experiment. The registrations remained relatively stable until they suddenly rose in the summer of 2015, shortly after the 1st wave of our field experiment.

The large and unexpected influx of refugees has put a lot of strain on the German asylum system, which was at that time only designed to accommodate about 150,000 asylum seekers per year (Schaub, Gereke and Baldassarri 2020: 691). Counties throughout Germany received more asylum seekers than they were prepared
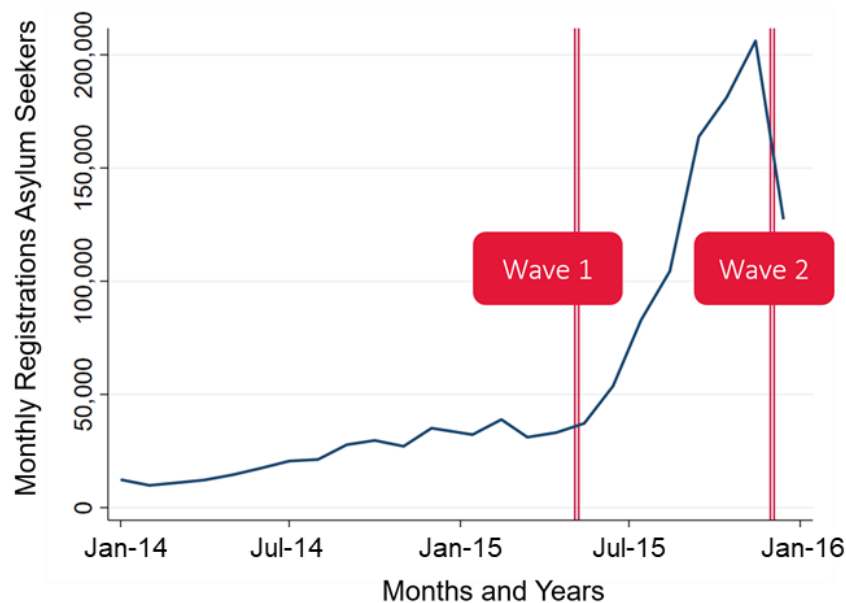
Figure 2: Monthly Number of Asylum Seekers and Timing of the Two Waves of our Field Experiment.
Note: This figure shows the number of newly registered asylum seekers in Germany from January 2014 to January 2016 and the timing of the two waves of our repeated field experiment on ethnic discrimination (each lasting one week). On average, about 1 refugee per 100 inhabitants arrived in the 401 counties in Germany between the dates of the two waves. Data: German Federal Ministry of the Interior and Community, own illustration.

to accommodate. In order to equalize the resulting burden on counties and municipalities, the geographical distribution of newly arriving refugees was organized according to a quota system ("Königsteiner Schlüssel"). Each federal state received refugees proportional to its population size (two-third weight), and to a lesser extent (one-third weight) its economic capacity, measured by tax revenues. Within the federal states, refugees were initially housed in central reception centers. Within three months, they were then moved to smaller refugee shelters in municipalities all over Germany, mostly based on similar quota systems at the level of counties. During the complete asylum process, all refugees were required by law to remain at their assigned place of residence. In 2015, asylum applications took roughly seven months to process (Schaub, Gereke and Baldassarri 2020). This means that almost all refugees were still living in their assigned municipality during the 2nd wave of our field experiment (for more details and statistics on the allocation process: Aksoy, Poutvaara and Schikora 2020).[6] Taken together, these administration procedures imply that per capita refugee immigration to counties was exogenous to most county characteristics that might affect housing discrimination, such as the size of previous immigrant populations. (The plausibility of this assumption will be discussed in more detail later.)

Data

We conducted a paired e-mail correspondence test (field experiment) with two waves in 2015. The first wave took place before the start of the increased refugee immigration from May 4 to May 8, and the second wave took place at the peak of the refugee immigration from November 30 to December 4. Correspondence tests are considered the gold standard for measuring discrimination (Baldassarri and Abascal 2017, Ross 2017). Two male applicants, one of whom signaled a German, the other a Turkish background, applied for the same rental housing unit.[7] Applicant ethnicity was indicated by 30 different typical Turkish and German names each. We chose a paired testing (within-subject) design with two applications per advertisement to give housing suppliers at least one viable alternative to the potentially disliked applicant and to take advantage of the higher power of a within-testing design (Charness, Gneezy and Kuhn 2012): The timing of the experiment, characteristics of the tested housing units/suppliers, and their regional context are constant for the pair of applications, which increases the statistical power to detect effects of ethnicity.

Besides ethnicity, we also varied the (amount of) information on applicants' family and labor market status. All applicant characteristics were fully crossed based on a *D*-efficient design, the gold standard for optimal orthogonal and balanced experimental designs (Auspurg and Hinz 2015). Although ethnicity was always varied between the two applications to the same housing unit (i.e., one application was always sent by a Turkish, one by a German applicant), the levels of other experimental factors could either be the same or different between the two applications to the housing unit. Such multifactorial designs prevent the stimulus of ethnicity from being confounded with the composition of the candidate pool applying to the same vacancy (Phillips 2019). Moreover, they allow for the standardization of key sociodemographic variables between the two groups of Turkish and German applicants, thus providing clearer evidence that ethnicity (and not, e.g., different financial backgrounds) causes potential discrimination. In addition, such designs help to conceal the nature of the experiment to housing suppliers avoiding problems of reactivity. The risk of detection was also minimized by using slightly different versions of salutations and other phrases for the two e-mail applications sent to one housing offer (different phrases were randomly assigned).

The paired e-mails were automatically sent with a time interval of about one hour, and the order of who applied first (the Turkish or the German applicant) was alternated to avoid confounding ethnicity with possible order effects. Following standards in such experiments, we relied on the observation of whether housing suppliers replied to one or both e-mails to identify unequal treatment (discrimination).[8] At the level of the tested housing units, this results in three possible outcomes:

a. Both applicants receive/do not receive a response: = equal treatment;

b. Only the German applicant receives a response: = discrimination against the Turk;

c. Only the Turkish applicant receives a response: = discrimination against the German.

Following standard approaches, we define the gross discrimination rate of Turks (Germans) as the percentage of cases with outcome *b* (*c*); and the net discrimination rate of Turks that indicates systematically greater discrimination against Turks compared to Germans as $b - c$. We will use the more "conservative" net discrimination rate (Ondrich, Ross and Yinger 2000) as the main outcome and use gross discrimination only in robustness analyses.

In both waves, we sampled 2,500 listings for rental housing units with 2 to 4 rooms to be tested in our experiment. A web-scraping procedure was used to randomly select these housing units from a major online housing platform in Germany. For ethical reasons and to follow standard procedures for field experiments in the housing market (Zschirnt 2019, Auspurg, Schneck and Thiel 2020), we sampled on the level of suppliers and not housing units, meaning we tested each supplier only once. Few housing units ($N = 188$) were excluded from the analysis sample as they were no longer available on the housing platform at the time the e-mail of the 2nd applicant was planned to be sent. In these cases, a paired test was not feasible. A few more units ($N = 13$) were excluded as no information on their regional location was available, making it impossible to measure moderator variables (i.e., the share of foreigners living in the region) or control variables (e.g., the federal state in which the housing unit was located). In total, the analysis sample included 4,799 rental units: 2,389 tested in the 1st and 2,410 tested in the 2nd wave.

Identification Strategy

The central theoretical estimand (cp. Lundberg, Johnson and Stewart 2021) of our study is the effect of refugee immigration in the second half of 2015 on discrimination of Turks in the rental housing market. The empirical estimate we use is the difference in discrimination found before (control group, 1st wave) and at the peak of refugee immigration (treatment group, 2nd wave). A simple estimate of the average treatment effect (ATE) is the mean difference θ between the level of net discrimination found in the 2nd versus 1st wave of our field experiment:

$$\theta = (b - c)_{wave2} - (b - c)_{wave1}$$

The unbiased identification of this ATE requires that differences in the level of discrimination between the two waves (i.e., control and treatment group) must solely be due to the strong influx of refugees during the refugee crisis. A key challenge to this assumption in our study is that we could not test the same housing units/suppliers in both waves but had to rely on different samples for practical and ethical reasons.[9] For a valid identification, it is then crucial that the samples observed in both waves are balanced in terms of (unobserved) covariates that could also influence the level of discrimination (Muñoz, Falcó-Gimeno and Hernández 2020). Because we used exactly the same design for the field experiment and sampling of housing units/suppliers for both waves, and the time interval between the two waves was short ($\sim$ 7 months), it seems plausible that most covariates are

balanced, which is supported by statistical tests for nearly all observable covariates (see *S3.2 online supplement*).[10]

 We nevertheless account for possible imbalances in two ways. First, we use multivariable regression models where we regress net discrimination on a wave dummy to estimate the treatment effect and a large bunch of covariates to account for possible changes in the composition of tested housing units/suppliers or housing markets (on the apartment level: number of rooms, rent per sqm, private housing supplier yes/no – on the county level: located in a metropolitan area yes/no, Gross Domestic Product (GDP) per person employed, population density, share of foreigners, unemployment rate, vacancy rates and voter share of green-party; see *S1.2 online supplement* for details on the regression approach). We also included federal state fixed effects that allow to net out all time-constant variables, observed or unobserved, on this level. Second, we repeat our robustness analyses for various subsamples, such as only private versus non-private housing suppliers, ensuring perfect balance of those main covariates (see *S3.1 online supplement*).

 To identify treatment heterogeneity by regions' prior exposure to immigrant populations (our second research goal), we compare differences in changes in discrimination across the two waves between counties with different sizes of foreign populations (measured by official per capita numbers in 2014, the year before the refugee crisis; in contrast to the newly arrived refugees, most of the 2014 foreigners already had a long-term residence permit for Germany, so their population share is a good indicator of the prior contact of residents with immigrants).[11] To estimate the possible moderation by this variable, we extended our multivariable regressions by an interaction term between the wave dummy and the size of foreign populations.

 Still, several assumptions must be met for a valid identification (Muñoz, Falcó-Gimeno and Hernández 2020). These are evaluated later in the section on threats to identification. More details on the research design, analytical strategy, and robustness of results are provided in our *online supplement*.

Results

Main Findings

In total, our pairwise-matched applications resulted in data on 9,598 rental requests: ∼2,500 housing units per wave, each contacted by one Turkish and one German applicant. First, we report the overall discrimination rates, then we continue with the changes across waves. Averaged over both waves of our experiment, the gross discrimination rates of Turks and Germans were 14.3 and 4.0 percent (see Table **??**). The statistically significant net discrimination rate (difference between the two gross discrimination rates) of about 10 percentage points matches well with the discrimination rates found in previous correspondence tests in Germany (Auspurg, Schneck and Hinz 2019).

 Applying the correlative cross-sectional approach used in previous literature, we find that gross and net discrimination against Turks were stronger in counties with a *lower* share of foreigners. This can be shown descriptively (see Figure 3) and in multivariate regression analyses that control for other contextual factors such as

Table 1: Total Discrimination Rates (Both Waves Pooled).

| | | Turkish Applicants | |
		Response	No response
German Applicants	Response	**Equal treatment** 49.2% N=2,360	**Gross discrimination Turkish** 14.3% N=686
	No response	**Gross discrimination German** 4.0% N=194	**Equal Treatment** 32.5% N=1,559

Notes: Net discrimination T: 14.3% - 4.0% = 10.3%

McNemar's χ^2 (1) = 275.1, p<0.001

The number of tested housing units was $N = 4,799$. The overall response rate was 58.3%. For the German (G) applicant the response rate was 63.5%, for the Turkish (T) applicant it was 53.2%.

regional vacancy rates or type of supplier (private supplier vs.commercial agency). Following these results, one might conclude that a larger immigrant population *lowers* the risk of discrimination.

However, to see whether there is a *causal* effect of immigration on discrimination, we have to draw on the exogenous variation in immigration during the refugee crisis. To identify the effect of the main treatment variable (i.e., refugee immigration), we first compare the discrimination rates from the 1st wave of the experiment (left panel in Figure 4) with those observed in the 2nd wave (right panel in Figure 4). There was a slight decrease in the gross discrimination rates of both Turks and Germans of around 1 percentage point each. This was due to a slight increase in both categories of equal treatment: Both applications received a response or did not receive a response slightly more often in the 2nd wave (details in *S2.1 online supplement*). This small parallel decrease in gross discrimination rates kept the net discrimination rate relatively stable at roughly 10 percentage points. Second, the finding that there was no credible evidence for a change in the level of net discrimination (and also gross discrimination rates) between waves is confirmed by multivariable regressions where we use a wave dummy (and in robustness analyses metric variables specifying the (relative) magnitude of refugee immigration to different counties) to identify the treatment effect (estimates shown in *S2.1 online supplement*, in particular Figure S2). The change in discrimination across waves was throughout very small in size and not statistically significant.

Heterogeneous Effects: Regions with Varying Levels of Previous Immigration

Some counties in Eastern Germany were not accustomed to foreigners (share of immigrants < 2 percent in 2014, whereas the mean share in Germany was 9.3 percent in this year) until refugees were allocated by law during the refugee crisis. In these regions, the influx of refugees may have been particularly salient and fueled anti-immigrant sentiments and inclinations to discriminate. In contrast, in some
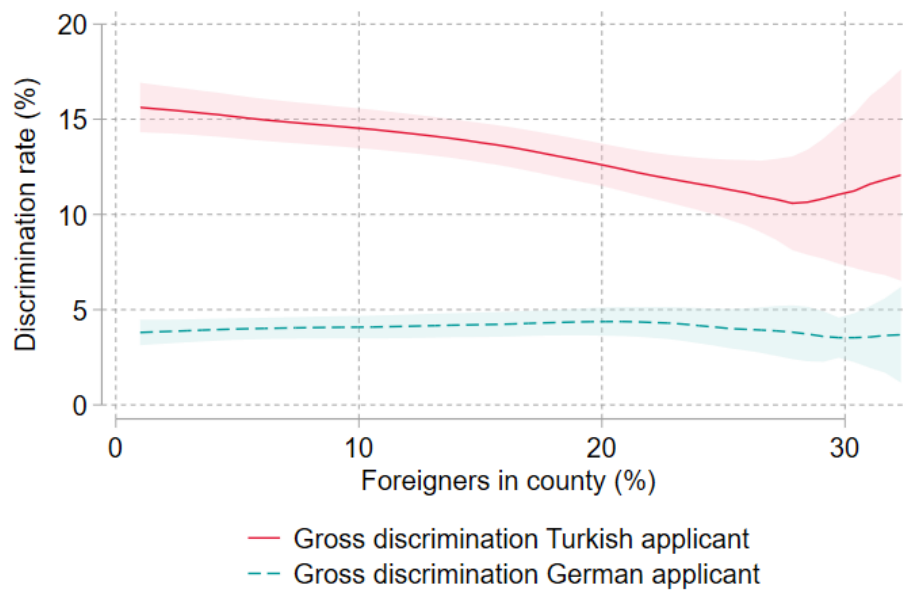
Figure 3: Overall Discrimination Rates by Share of Foreigners in a County.
Note: This figure shows a nonparametric polynomial smoothing of the association of the gross discrimination of Turks (red, solid line) and Germans (green, dashed line) over counties with different shares of foreigners (population with no German nationality, according to official statistics) together with 95 percent confidence intervals. The gap between both discrimination rates is the net discrimination rate. One can see that the gross discrimination of Turks and accordingly the net discrimination rate decline with increasing shares of foreigners living in a county. The figure is based on N = 4,799 tested apartments in both waves and was produced with the Stata command lpolyci. In order to identify stable trends behind idiosyncratic (outlier) effects, a wide bandwidth was used (6pp).

urban regions in Western Germany, at least one-third of the residents had foreign citizenship prior to the refugee crisis. Further immigration during the refugee crisis has made little quantitative difference in these regions and may therefore have been perceived as less disruptive (Hopkins 2010).

Figure 5 summarizes the net discrimination rates in counties with varying shares of foreigners prior to the refugee crisis: from the lowest (very low: 1 percent-6 percent) to the highest quartile (very high: 14 percent - 32 percent). For easier interpretation, we only show net discrimination rates (detailed results, including gross discrimination rates, are provided in *S2.2 online supplement*, Table S2). Overall, net discrimination was lower in counties with a higher percentage of foreigners (panels on the right). Although there were some changes in net discrimination rates in some quartiles (e.g., the net discrimination in the second quartile decreased from 14 to 11 percentage points across waves), there was no clear pattern that the change in discrimination was moderated by the size of the foreign population). In all panels, there is a large overlap in the confidence intervals for the net discrimination rates observed in both waves. We therefore conclude that the main finding that immigration has no substantial effect on discrimination also appears to hold for different levels of prior exposure to immigration.
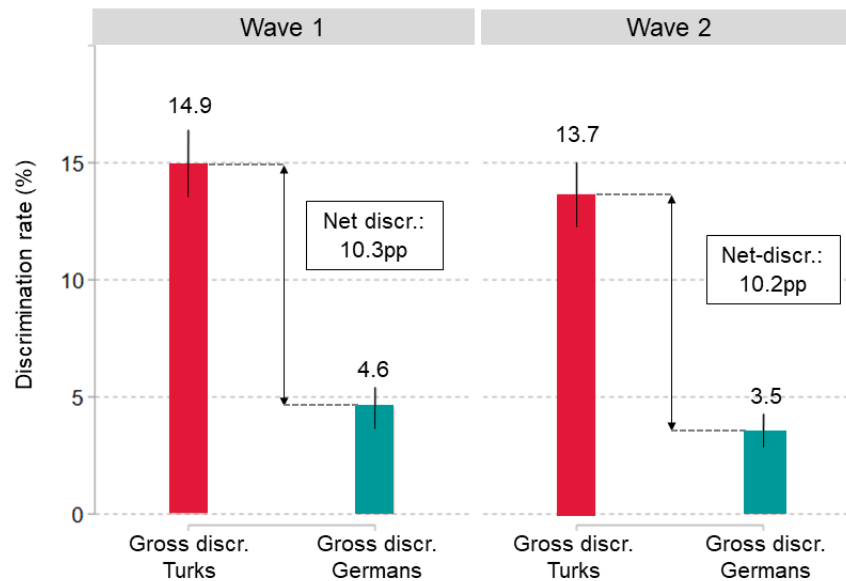
Figure 4: Gross and Net Discrimination Rates by Wave.
Note: The bars show the gross discrimination rates in percent. The net discrimination, which is the difference between the gross discrimination rate of Turks and Germans, is indicated in percentage points (pp). The sample comprises 2,389 tested housing units in the 1st wave and 2,410 housing units in the 2nd wave.

Threats to Identification and Robustness Checks

The benefit of our identification strategy is that we rely with the refugee crisis on a kind of exogenous shock. The estimation of an unbiased causal effect using such "unexpected event" (Muñoz, Falcó-Gimeno and Hernández 2020) during the field phase of an individual data collection relies on assumptions about "excludability" and "ignorability," which we shortly discuss in the following (mainly based on Muñoz, Falcó-Gimeno and Hernández 2020, Harding and Nwokolo 2023).[12] The *excludability* assumption is that identified differences in discrimination levels between the two waves (control and treatment group) are solely due to refugee immigration. This assumption could be violated in several ways. Differences in the level of discrimination could also be caused (or counteracted) by simultaneous events happening at the same time as the refugee immigration, by unrelated preexisting time trends, or seasonal effects that may also have impacted the level of ethnic discrimination. Events or trends that *equally* affected housing suppliers' responses to Turkish and German applicants, such as generally responding to e-mail inquiries more frequently in the winter compared to the spring, would not bias the unequal treatment measured by the net discrimination rates. It seems plausible that most trends or seasonal/periodic effects would have such consistent effects on both applicants. However, some intervening shocks or trends might have affected only discrimination against Turkish applicants, such as the Islamist terrorist attacks in Paris that occurred shortly before the 2nd experiment.
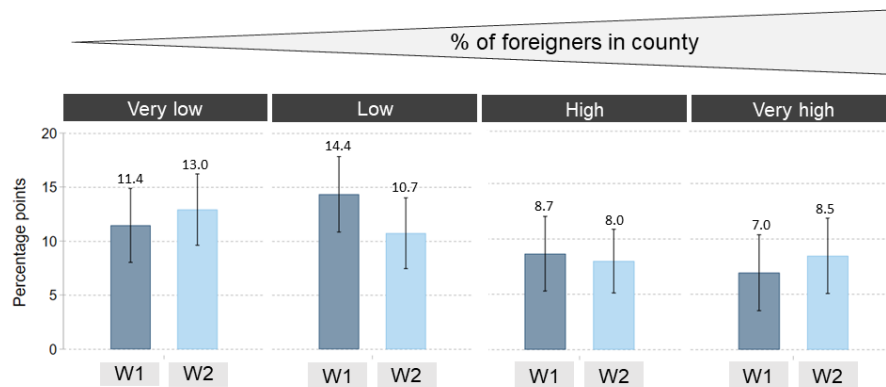
Figure 5: Net Discrimination Rate by Level of Foreigners and by Wave.
Note: This figure shows discrimination rates in percentage points together with 95 percent confidence intervals. W1 (W2) shows results for the 1st (2nd) wave. The level of foreigners in a county (per 100 inhabitants) is split into quartiles (very low: 1.0%- 6.0 %, low: 6.1%-9.8 %, high: 9.9%-14.3%, very high: 14.4%- 32.3%). At least 472 tested housing units per quartile and wave.

To circumvent such confounding factors, we also use a metric measure of refugee immigration: the magnitude of immigration to different counties, which ranged from 0.3 to 1.2 refugees per 100 inhabitants ("outlier" counties with exceptionally large reception centers excluded). We see no reason that this exogenous variation in the size of immigration would be systematically correlated with temporal trends or seasonal/periodic effects, at least when we control for tax revenues (measured by the GDP proxy). This was supported by balance checks (see *S3.2 online supplement*). We are not aware of any other historical events or trends in the second half of 2015 that may have had an impact on ethnic discrimination in Germany or even remotely similar effects on public attention or population composition. With respect to general time trends, there is evidence that ethnic discrimination in the housing market is decreasing in Europe, but at such a slow pace that no substantial change can be expected within 7 months.[13] Finally, differences in discrimination levels may have been caused by collateral events triggered by the refugee crisis, such as the strong media discussions, and not the refugee immigration in itself. We admit that we cannot separate these joint events; but do not see this as a major problem, because these collateral events are also in the focus of theories on group threats and conflicts.

The second key assumption for identification is *ignorability*: For the tested housing units/suppliers, the assignment to treatment or control group must be independent of the potential outcome of discrimination, that is, the assignment must be as good as random (Muñoz, Falcó-Gimeno and Hernández 2020). Again, several threats to this assumption exist. First, the samples used in both waves (i.e., the treatment and control groups) could differ due to an imbalance on observables, which may be related to the outcomes of interest. We deal with potential problems arising from the imbalance on observable covariates by controlling for the covariates noted in the method section or splitting the sample by covariates, giving *conditional ignorability*. Another potential threat stems from unobserved confounders. By

including fixed effects for federal states, we correct for possible imbalance in the allocation of housing units/suppliers across federal states and achieve balance even in unobserved time-constant confounders on the federal level.

For an unbiased identification of treatment heterogeneity (our second research goal), also the distribution of refugees to counties with different previous exposure to immigration should be as good as random; otherwise, we might confound the interaction of the wave dummy with the size of foreign populations we are interested in with a different strength of treatment (i.e., different size of refugee immigration). The setting of the refugee crisis with a quasi-random distribution of refugees across regions due to the used quota design makes us confident that we could reach at least *conditional ignorability* (see also Aksoy, Poutvaara and Schikora 2020 for tests of exogeneity assumptions in the distribution of refugees across counties). In addition, using the size of refugee immigration as an alternative treatment in robustness analyses serves as a further test for the plausibility of this assumption. In further robustness checks, we also used the relative change in refugee immigration.

For the unbiased identification of the ATE or treatment heterogeneity, it is in addition crucial that the event of the refugee crisis was unpredictable for housing suppliers. (Otherwise, suppliers may have already adjusted their behavior in the 1st wave in anticipation of the crisis, which would hinder a clear separation of treatment and control group.) This assumption is certainly met here: In May 2015, the suspension of the Dublin Agreement and accordingly the strong refugee movements were not foreseeable.

Finally, threats to identification might result from an invalid measurement of discrimination. A drawback of experimental within-designs is that housing suppliers might detect them. We have taken all kinds of precautions, for example, varying more information than just ethnicity.[14] Only one out of our 5,000 sampled housing units showed evidence that our experiment was detected.[15] There might have been further cases we did not notice. This would have led to an underestimation of discrimination due to possible demand effects (Charness, Gneezy and Kuhn 2012). However, as long such bias is not tied to one of the two waves of our experiment, this would not have biased the estimation of the effect of refugee immigration. We have no reason to assume that detection was more likely in the 2nd wave, because we did not publish any information on the field experiment until now. The other known disadvantage of within-designs, order effects, was eliminated by balancing the order in which the two applications were sent.

We also tested the robustness of our results with other statistical models, including instrumental variable approaches (see *S3.1 online supplement* for all robustness and *S3.2* for balance checks). In these robustness analyses, we included nonparametric methods, which do not require assumptions about the functional form and thus allow the identification of possible nonlinear effects or "tipping points" (Galster 2014). As further alternative treatments, we used the number of male refugees in a county. None of these analyses changed our core conclusion: Refugee immigration in the course of the European refugee crisis in 2015 had no meaningful impact on the level of discrimination of Turks observed in our experiments. The conclusion that discrimination rates did not change significantly across waves was also con-

firmed by using alternative outcomes to measure discrimination, such as explicit invitations to visit a housing unit or response times to the initial request. To test whether changes in discrimination occur on a more fine-grained spatial level, we conducted additional analysis at the neighborhood level using walking distances to refugee shelters as an alternative treatment variable. Again, we uniformly found no significant change in discrimination rates. Therefore, we are confident that we have not missed more subtle or regional changes in discrimination levels.

Discussion and Conclusion

In this study, we analyzed the effect of a strong and unexpected mass immigration on ethnic discrimination: We studied the effects of immigration of refugees from mostly Muslim countries on the discrimination of Turks in the German rental housing market.

A particular strength of our study is that we could leverage the mostly exogenous assignment of refugees to municipalities in Germany, which helps to overcome confounder and self-selection bias that may have biased previous research. The main result is surprising: Although Germany has experienced increasing concerns about immigration (Torres 2022) since the refugee crisis, as well as decreasing willingness to host some immigrant groups (Czymara and Schmidt-Catran 2017), hostile attacks against refugees (Jäckle and König 2018, Wagner et al. 2020), and increasing supporters of right-wing parties opposing immigration (Dostal 2017), we have not seen an increase in discrimination against Turks in the rental housing market in Germany. Our results are robust across regions accustomed to varying levels of immigration prior to the crisis. They are in line with Schaub, Gereke and Baldassarri (2020), who found no evidence for an effect of refugee immigration in Germany on ethnic/religious discrimination measured in a lab experiment in 2018. Our results allow to strengthen the external validity and draw conclusions about a natural setting, the German housing market.

How can we explain this result? A first explanation would be that no spillover effects occur from refugees emigrating from the Middle East or Africa to the Turks tested in our experiment. This could theoretically be the case if natives do not lump all immigrants with a supposed Muslim or Middle Eastern background together but instead draw finer boundary lines between immigrating refugees and Turkish immigrants who have lived in Germany for a longer time. This seems, however, implausible in the light of previous research that reported strong spillover effects in the form of the refugee crisis fueling *general* anti-Muslim sentiments, including negative sentiments against Muslim immigrants that lived in the host countries for generations (Dinas et al. 2019).[16]

A second reason could be that housing providers are less inclined to develop anti-Muslim sentiments, for example, because property owners often have a high level of education and per se wealth, which better protects them from competition with low-status immigrants. Evidence shows that majorities with higher socioeconomic backgrounds displayed less intolerance toward minorities before and after the European refugee crisis (Kromczyk, Khattab and Abbas 2021). So far, however, we can only speculate about the extent to which this translates to property owners

in the housing market. To date, only cross-sectional evidence relates aggregate measures of anti-immigrant attitudes to regional levels of discrimination (see, e.g., Lacroix, Ruedin and Zschirnt 2022). This approach is, however, prone to the risk of ecological fallacy and confounder bias. Future research should attempt to measure attitudes and feelings of threat at the micro level of those involved in the rental processes.

A third explanation would be that—at least in the short run—anti-migration attitudes do not necessarily translate into discrimination. Prejudiced and/or threatened property owners and agencies may be unwilling to bear the costs of animus discrimination. Profit-seeking could be a more important driver of discrimination than negative feelings toward immigrants. This aligns with theories on statistical or customer discrimination (Arrow 1971 , Phelps 1972, for evidence on the housing market see, e.g., Ewens, Tomlin and Wang 2014), but would still require future research that analyzes motives for discrimination in more detail. In any case, our results suggest that researchers interested in hostile acts should also measure behavioral outcomes, not just attitudes.

In contrast to fully randomized experiments, natural experiments always rely on quasi-randomization, which may threaten internal validity. Although we tried to evaluate the underlying assumptions of the design used as well as possible, some threats, such as unobserved confounders, might still exist. It is also important to keep in mind our limited target population, suppliers on the German online rental housing market. Ethnic discrimination in the housing market seems to be more pronounced in Germany than in other European countries, and Turks seem to be particularly affected (Auspurg, Schneck and Hinz 2019, for cross-country evidence on the labor market see Quillian et al. 2019, Flage 2018). Thus, we examined a "more likely" case in which (an increase in) discrimination could be expected. Nevertheless, we cannot rule out the possibility that scope conditions for discrimination differ by country or market and that there are effects of immigration on discrimination elsewhere. For these reasons, further studies on other countries and markets, such as the labor or consumer market, are desirable, as well as studies on discrimination in "everyday encounters" (Zhang, Gereke and Baldassarri 2022).

Our study has implications for immigration policies and points out further fruitful avenues for future research. First, many countries are currently experimenting with refugee allocation procedures. Some, such as Canada, allow refugees to move freely within the country and try to channel them, if at all, to regions with good labor market options (Bansak et al. 2018). Similarly, in the United States, refugees are not required to stay in an assigned place of residence (Bruno 2017). Studies examining the economic integration of refugees and the economic well-being of host communities have consistently found that both are optimized when refugees are settled in more prosperous regions that offer them better labor market options (Martén, Hainmueller and Hangartner 2019). However, these policies lead to a strong regional concentration of refugees in a few metropolitan areas (Rose 2019). In contrast, several European countries, such as Germany and Switzerland, aim to distribute refugees homogeneously across the country to equalize the burden on the welfare funds of single municipalities and to avoid the formation of ethnic enclaves and housing shortages. Our experiment allowed us to evaluate this policy

in Germany regarding potential unintended side effects in the form of increased ethnic housing discrimination of earlier immigrants. Discrimination was found to be *generally* lower in counties with many immigrants and thus higher ethnic diversity. According to our results, this correlation was most likely not caused by the size of immigrant populations. Instead, we suspect that confounding spatial factors and previous immigrants' self-selection into specific regions are responsible for this correlation. Even without diving deeper into the underlying mechanisms (e.g., whether this represents a general composition effect or whether it is caused by more immigrant housing suppliers in these counties) we can now conclude that allowing more refugee migration to these ethnically diverse regions does not seem to have negative side effects in terms of increased ethnic housing discrimination against Turks. On the contrary, our results suggest that migrants can expect less discrimination in ethnically mixed regions, which should improve their chances of integration. However, more research is needed on the long-term trend in discrimination,[17] as well as further insights about additional migrant groups (including possible discrimination against refugees).

Beyond these substantial findings, our combination of field and natural experiment is an interesting case study grasping the essence of discrimination research: causal research on the "why," "when," and "where" questions of discrimination. Studies using context (i.e., moderator) variables and other conditions for discrimination that are not randomly assigned suffer from post-treatment bias (Montgomery, Nyhan and Torres 2018). Thus, more studies with random variation in such variables are necessary. Natural experiments may help achieve such variation. For example, events that are already known to alter attitudes such as terrorist attacks (see, e.g., Legewie 2013), may serve as treatment or instrumental variables. One key recommendation is to design more longitudinal field experiments that are "always-on" (i.e., in the field), or at least have several (long) field periods to increase the likelihood of capturing natural experiments by chance.

Data Note

We used data we collected in the research project "Ethnic Discrimination and Segregation in German Housing Markets" funded by the Wolfgang and Anita Bürkle Foundation. Our replication files (data and Stata do-files) are available on the OSF platform (https://doi.org/10.17605/OSF.IO/S8R2H).

Notes

1 Home ownership is much more common among natives than among migrants in Germany. Official statistics for 2014 show, for example, that people without a migration background were much less likely to live in residential property (54.8 percent) than people with a migration background (34.5 percent, German Federal Statistical Office - Destatis 2017). Migrants being much more dependent on the rental housing market lends particular importance to studies of ethnic discrimination in this market.

2 Contact theories assume that larger immigrant populations may increase the chances of positive contact (e.g. building friendships) under certain conditions. Positive contact

reduces hostility and discrimination by promoting empathy and understanding (Allport 1954, Elwert, Keller and Kotsadam 2023, Pettigrew and Tropp 2006). According to Allport's original specification, intergroup contact reduces prejudice primarily when natives and immigrants have the same status, and when they cooperate and work toward a shared goal. Although later theories suggests that these conditions are facilitative rather than essential (Pettigrew & Tropp, 2006), one scope condition has crystallized as particularly important: Contact must provide opportunities to form friendships (c.f. Bohman and Miklikowska 2021). However, during the first phase of refugee integration in which our experiment took place, contacts with refugees were mostly fleeting, non-repetitive encounters that lack the depths of contact for empathy to evolve (see e.g. the survey results of Schmidt, Weick and Gloris 2019 on contact frequencies).

3 Regions as units of comparisons reached from neighborhoods within a single city (e.g. Auspurg, Hinz and Schmid 2017) over different districts within countries (Van der Bracht, Coenen and Van de Putte 2015) up to comparing different countries (Hexel et al. 2019). Some studies also analyzed associations between the level of ethnic/racial segregation or ethnic diversity with the level of ethnic/racial discrimination (for an overview: Krysan and Crowder 2017: Ch. 9).

4 According to the Dublin III Regulation, refugees were only eligible to apply for asylum in the country where they had first entered the European Union.

5 In 2015, immigration to Germany was higher than ever before, according to statistics from the German Federal Statistical Office - Destatis (2016). 2.16 million people moved to Germany, of which about 0.9 million were asylum seekers who fled from (civil) wars and poverty, mainly in Syria, Afghanistan, Iraq, and some south-eastern European countries (German Federal Ministry of the Interior and Community - BMI 2016). In 2015, about 7.8 Mio inhabitants in Germany (10 percent) did not have a German citizenship, and about 17.1 Mio (21 percent) had a migration background (with the largest immigrant groups being Turks, followed by Poles and immigrants from the Russian Federation).

6 Even between 2016 and 2018, only 8 percent of the refugees in Germany moved to another federal state (Aksoy, Poutvaara and Schikora 2020). These authors, who tested a variety of variables, found that county-level population size was the only statistically significant predictor of the number of asylum seekers assigned per county.

7 Turks formed the largest ethnic minority in Germany in 2015. We focused on this group in our experiments because there was already evidence that Turkish immigrants in Germany were particularly affected by negative prejudice and discrimination (in the housing market). One reason for this discrimination is probably their presumed Muslim religion (Auspurg, Hinz and Schmid 2017, Auspurg, Schneck and Hinz 2019).

8 Other outcomes (the content of responses and response times) are used in robustness checks.

9 Rental housing units are typically advertised only few days or weeks, meaning that most housing units/suppliers tested for discrimination in our field experiment were available only in one of the two waves. Testing only suppliers available at both waves would have strongly impacted the external validity of our study. As mentioned above, we also decided to test each supplier only once for ethical reasons.

10 The only two exceptions were a small increase in the proportion of private suppliers (that tend to discriminate more likely compared to real-estate agencies, see Flage 2018) due to a legislative reform, and a slight decrease in the share of housing units located in counties with a high share of voters for a left-wing party (the Green party) supporting immigration in the last federal election (2013). When not accounted for, both imbalances may lead to a slight overestimation of the increase in discrimination across waves.

11 Foreigners are individuals without a German nationality. For regional units (counties) there are official statistics on a yearly basis only for foreigners, not the whole population with a migrant background. However, both variables are strongly correlated (see S1.1 online supplement). Germany is divided into 402 counties (varying in size from 34,260 to 352,0031 inhabitants in 2015), composed of a city and/or a rural region. Overall, the field experiment was run in all federal states and in 328 of the 401 counties in both waves.

12 Those designs are typically used for events happening during the field phase of a public opinion survey (Muñoz, Falcó-Gimeno and Hernández 2020). We transfer the design to field experiments. Compared to the designs relying on surveys, our setup based on a field experiment allows to preclude threats to identification that result when individuals self-select into a survey study.

13 A particularly good indication of this is that the level of net discrimination measured in our study is about the same as what was measured in similar field experiments in Germany in 2011 and 2012 (Schmid 2015, Auspurg, Hinz and Schmid 2017).

14 Detection is also presumably much less likely in experiments on the housing market than in experiments on the labor market (Weichselbaumer 2015) because a much less formalized application process (e.g., without certificates) is used.

15 This case was excluded from our sample as one case where the apartment was not actually available to the second applicant.

16 Including refugees as another group in our experiments would have allowed for more accurate conclusions about possible spillover effects, as such a design would allow us to compare treatment effects for refugees and Turks. However, as with many natural experiments, we could not anticipate the refugee crisis and therefore did not include refugees as an ethnic group in the 1st wave of our experiments.

17 Research on long-term effects would be very valuable because there are also some theoretical arguments that stocks of immigrants may have different effects on attitudes towards immigration and dealing with immigrants than flows that we observed in our study (Brady and Finnigan 2014).

References

Aksoy, Cevat, Panu Poutvaara and Felicitas Schikora. 2020. "First Time Around: Local Conditions and Multi-dimensional Integration of Refugees." *IZA Discussion Paper No. 13914.* https://doi.org/10.31235/osf.io/nsr8q

Alba, Richard. 2005. "Bright vs. Blurred Boundaries: Second-Generation Assimilation and Exclusion in France, Germany, and the United States." *Ethnic and Racial Studies* 28(1):20-49. https://doi.org/10.1080/0141987042000280003

Allport, Gordon W. 1954. *The Nature of Prejudice*. Reading, MA: Addison-Wesley.

Arrow, Kenneth J. 1971. "The Theory of Discrimination." *Princeton University. Industrial Relations Section Working Paper No. 30A.*

Auspurg, Katrin and Thomas Hinz. 2015. *Factorial Survey Experiments*. Thousand Oaks: SAGE.

Auspurg, Katrin, Thomas Hinz and Laura Schmid. 2017. "Contexts and Conditions of Ethnic Discrimination: Evidence From a Field Experiment in a German Housing Market." *Journal of Housing Economics* 35:26-36. https://doi.org/10.1016/j.jhe.2017.01.003

Auspurg, Katrin, Andreas Schneck and Thomas Hinz. 2019. "Closed Doors Everywhere? A Meta-Analysis of Field Experiments on Ethnic Discrimination in Rental Housing

Markets." *Journal of Ethnic and Migration Studies* 45(1):95-114. https://doi.org/10.1080/1369183x.2018.1489223

Auspurg, Katrin, Andreas Schneck and Fabian Thiel. 2020. "Different Samples, Different Results? How Sampling Techniques Affect the Results of Field Experiments on Ethnic Discrimination." *Research in Social Stratification and Mobility* 65:100444. https://doi.org/10.1016/j.rssm.2019.100444

Baldassarri, Delia and Maria Abascal. 2017. "Field Experiments Across the Social Sciences." *Annual Review of Sociology* 43(1):41-73. https://doi.org/10.1146/annurev-soc-073014-112445

Baldini, Massimo and Marta Federici. 2011. "Ethnic Disrimination in the Italian Rental Housing Market." *Journal of Housing Economics* 20(1):1-14. https://doi.org/10.1016/j.jhe.2011.02.003

Bansak, Kirk, Jeremy Ferwerda, Jens Hainmueller, Andrea Dillon, Dominik Hangartner, D. Lawrence and J. Weinstein. 2018. "Improving Refugee Integration Through Data-Driven Algorithmic Assignment." *Science* 359:325-29. https://doi.org/10.1126/science.aao4408

Becker, Gary S. 1957. *The Economics of Discrimination*. Chicago: University of Chicago Press.

Betts, Julian R. and Robert W. Fairlie. 2003. "Does Immigration Induce 'Native Flight' From Public Schools Into Private Schools?" *Journal of Public Economics* 87(5):987-1012. https://doi.org/10.1016/s0047-2727(01)00164-5

Blalock, Hubert M. 1967. *Toward a Theory of Minority-Group Relations*. New York: Wiley.

Blumer, Herbert. 1958. "Race Prejudice as a Sense of Group Position." *The Pacific Sociological Review* 1(1):3-7. https://doi.org/10.2307/1388607

Böhm, Robert, Maik M. P. Theelen, Hannes Rusch and Paul A. M. Van Lange. 2018. "Costs, Needs, and Integration Efforts Shape Helping Behavior Toward Refugees." *Proceedings of the National Academy of Sciences* 115(28):7284. https://doi.org/10.1073/pnas.1805601115

Bohman, Andrea and Marta Miklikowska. 2021. "Does Classroom Diversity Improve Intergroup Relations? Short- And Long-Term Effects of Classroom Diversity for Cross-Ethnic Friendships and Anti-immigrant Attitudes in Adolescence." *Group Processes & Intergroup Relations* 24(8):1372-90. https://doi.org/10.1177/1368430220941592

Brady, David and Ryan Finnigan. 2014. "Does Immigration Undermine Public Support for Social Policy?" *American Sociological Review* 79(1):17-42. https://doi.org/10.1177/0003122413513022

Bruno, Andorra. 2017. "Reception and Placement of Refugees in the United States." *Congressional Research Service Report R44878.*

Carlsson, Magnus and Stefan Eriksson. 2014. "Discrimination in the Rental Market for Apartments." *Journal of Housing Economics* 23:41-54. https://doi.org/10.1016/j.jhe.2013.11.004

Cea D'Ancona, M. Ángeles. 2018. "What Determines the Rejection of Immigrants Through an Integrative Model." *Social Science Research* 74:1-15. https://doi.org/10.1016/j.ssresearch.2018.05.008

Charness, Gary, Uri Gneezy and Michael A. Kuhn. 2012. "Experimental Methods: Between-Subject and Within-Subject Design." *Journal of Economic Behavior & Organization* 81(1):1-8. https://doi.org/10.1016/j.jebo.2011.08.009

Czymara, Christian S. and Alexander W. Schmidt-Catran. 2017. "Refugees Unwelcome? Changes in the Public Acceptance of Immigrants and Refugees in Germany in the Course

of Europe's 'Immigration Crisis'." *European Sociological Review* 33(6):735-51. https://doi.org/10.1093/esr/jcx071

Czymara, Christian S. 2021. "Attitudes toward Refugees in Contemporary Europe: A Longitudinal Perspective on Cross-National Differences." *Social Forces* 99(3):1306-33. https://doi.org/10.1093/sf/soaa055

Deole, Sumit S. and Yue Huang. 2020. "How Do New Immigration Flows Affect Existing Immigrants? Evidence From the Refugee Crisis in Germany." *GLO Discussion Paper, No. 579*.

Di Stasio, Valentina, Bram Lancee, Susanne Veit and Ruta Yemane. 2021. "Muslim by Default or Religious Discrimination? Results From a Cross-National Field Experiment on Hiring Discrimination." *Journal of Ethnic and Migration Studies* 47(6):1305-26. https://doi.org/10.1080/1369183x.2019.1622826

Dinas, Elias, Konstantinos Matakos, Dimitrios Xefteris and Dominik Hangartner. 2019. "Waking up the Golden Dawn: Does Exposure to the Refugee Crisis Increase Support for Extreme-Right Parties?" *Political Analysis* 27(2):244-54. https://doi.org/10.1017/pan.2018.48

Dostal, Jörg Michael. 2017. "The German Federal Election of 2017: How the Wedge Issue of Refugees and Migration Took the Shine off Chancellor Merkel and Transformed the Party System." *The Political Quarterly* 88(4):589-602. https://doi.org/10.1111/1467-923x.12445

Elwert, Felix, Tamás Keller and Andreas Kotsadam. 2023. "Rearranging the Desk Chairs: A large randomized field experiment on the effects of close contact on interethnic relations." *American Journal of Sociology*. https://doi.org/10.1086/724865

Ewens, Michael, Bryan Tomlin and Liang Choon Wang. 2014. "Statistical Discrimination or Prejudice? A Large Sample Field Experiment." *Review of Economics and Statistics* 96(1):119-34. https://doi.org/10.1162/rest_a_00365

Farley, Reynolds, Charlotte Steeh, Maria Krysan, Tara Jackson and Keith Reeves. 1994. "Stereotypes and Segregation: Neighborhoods in the Detroit Area." *American Journal of Sociology* 100(3):750-80. https://doi.org/10.1086/230580

Flage, Alexandre. 2018. "Ethnic and Gender Discrimination in the Rental Housing Market: Evidence from a Meta-Analysis of Correspondence Tests, 2006-2017." *Journal of Housing Economics* 41:251-73. https://doi.org/10.1016/j.jhe.2018.07.003

Fouka, Vasiliki and Marco Tabellini. 2021. "Changing In-Group Boundaries: The Effect of Immigration on Race Relations in the United States." *American Political Science Review* 116(3):1-17. https://doi.org/10.1017/s0003055421001350

Galster, George. 2014. "Nonlinear and Threshold Aspects of Neighborhood Effects." *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 66(1):117-33. https://doi.org/10.1007/s11577-014-0268-3

Gereke, Johanna, Joshua Hellyer, Jan Behnert, Saskia Exner, Alexander Herbel, Felix Jäger, Daria Szafran and Jule Schabinger. 2022. "Demographic Change and Group Boundaries in Germany: The Effect of Projected Demographic Decline on Perceptions of Who Has a Migration Background." *Sociological Science* 9:206-20. https://doi.org/10.15195/v9.a9

German Federal Ministry of the Interior and Community - BMI. 2016. "890,000 Asylum seekers in 2015." in *Press Release No. from September, 9, 2016*. https://www.bmi.bund.de/SharedDocs/pressemitteilungen/DE/2016/09/asylsuchende-2015.html (last access: 16.12.2022).

German Federal Office for Migration and Refugees - BAMF. 2016. *Migration Report 2015. Central Conclusions*. Nürnberg: Forschungszentrum Migration, Integration und Asyl.

German Federal Statistical Office - Destatis. 2016. "2015: Highest Levels of immigration and migration surplus in Germany." in *Press Release No. 246 from July, 14, 2016* https://www.destatis.de/DE/Presse/Pressemitteilungen/2016/07/PD16_246_12421.html (last access 16.12.2022).

German Federal Statistical Office - Destatis. 2017. "Marked Differences in the Housing Situation of People With and Without a Migration Background." in *Press Release No. 009/17 from January, 10, 2017* https://www.destatis.de/DE/Presse/Pressemitteilungen/2017/01/PD17_009_122pdf.pdf?__blob=publicationFile (last access 16.12.2022).

Ha, Shang E. 2010. "The Consequences of Multiracial Contexts on Public Attitudes toward Immigration." *Political Research Quarterly* 63(1):29-42. https://doi.org/10.1177/1065912908325255

Hainmueller, Jens and Daniel J. Hopkins. 2014. "Public Attitudes Toward Immigration." *Annual Review of Political Science* 17(1):225-49. https://doi.org/10.1146/annurev-polisci-102512-194818

Hangartner, Dominik, Elias Dinas, Moritz Marbach, Konstantinos Matakos and Dimitrios Xefteris. 2019. "Does Exposure to the Refugee Crisis Make Natives More Hostile?". *American Political Science Review* 113(2):442-55. https://doi.org/10.1017/s0003055418000813

Harding, Robin and Arinze Nwokolo. 2023. "Terrorism, Trust, and Identity: Evidence from a Natural Experiment in Nigeria." *American Journal of Political Science* Online First. https://doi.org/10.1111/ajps.12769

Hennig, Jakob. 2021. "Neighborhood Quality and Opposition to Immigration: Evidence from German Refugee Shelters." *Journal of Development Economics* 150:102604. https://doi.org/10.1016/j.jdeveco.2020.102604

Hexel, Ole, Lincoln Quillian, Anthony Heath, Devah Pager, Arnfinn Midtbøen and Fenella Fleischmann. 2019. "Do Some Countries Discriminate More than Others? Evidence from 97 Field Experiments of Racial Discrimination in Hiring." *Sociological Science* 6. https://doi.org/10.15195/v6.a18

Hopkins, Daniel J. 2010. "Politicized Places: Explaining Where and When Immigrants Provoke Local Opposition." *American Political Science Review* 104(1):40-60. https://doi.org/10.1017/s0003055409990360

Jäckle, Sebastian and Pascal D. König. 2018. "Threatening Events and Anti-Refugee Violence: An Empirical Analysis in the Wake of the Refugee Crisis during the Years 2015 and 2016 in Germany." *European Sociological Review* 34(6):728-43. https://doi.org/10.1093/esr/jcy038

Kaczan, David J. and Jennifer Orgill-Meyer. 2020. "The Impact of Climate Change on Migration: A Synthesis of Recent Empirical Insights." *Climatic Change* 158(3):281-300. https://doi.org/10.1007/s10584-019-02560-0

Korteweg, Anna and Gökçe Yurdakul. 2009. "Islam, Gender, and Immigrant Integration: Boundary Drawing in Discourses on Honour Killing in the Netherlands and Germany." *Ethnic and Racial Studies* 32(2):218-38. https://doi.org/10.1080/01419870802065218

Koubi, Vally. 2019. "Climate Change and Conflict." *Annual Review of Political Science* 22(1):343-60.

Kromczyk, Marcin, Nabil Khattab and Tahir Abbas. 2021. "The limits of tolerance: before and after Brexit and the German Refugee Crisis." *Ethnic and Racial Studies* 44(16):170-93. https://doi.org/10.1080/01419870.2021.1937670

Krysan, Maria and Kyle Crowder. 2017. *Cycle of Segregation - Social Processes and Residential Stratification*. New York: Russell Sage Foundation.

Kürschner Rauck, Kathleen. 2020. "Not in my Backyard! The 2015 Refugee Crisis in Germany." *University of St.Gallen, School of Finance Research Paper no 2020/04*. https://doi.org/10.2139/ssrn.3697189

Lacroix, Julie, Didier Ruedin and Eva Zschirnt. 2022. "Discrimination Driven by Variation in Social and Economic Conservatism: Evidence From a Nationwide Field Experiment." *European Sociological Review* 39(3):464-478. https://doi.org/10.1093/esr/jcac051

Legewie, Joscha. 2013. "Terrorist Events and Attitudes toward Immigrants: A Natural Experiment." *American Journal of Sociology* 118(5):1199-245. https://doi.org/10.1086/669605

Liebe, Ulf, Jürgen Meyerhoff, Maarten Kroesen, Caspar Chorus and Klaus Glenk. 2018. "From Welcome Culture to Welcome Limits? Uncovering Preference Changes Over Time for Sheltering Refugees in Germany." *PLOS ONE* 13(8):e0199923. https://doi.org/10.1371/journal.pone.0199923

Lundberg, Ian, Rebecca Johnson and Brandon M. Stewart. 2021. "What Is Your Estimand? Defining the Target Quantity Connects Statistical Evidence to Theory." *American Sociological Review* 86(3):532-65. https://doi.org/10.1177/00031224211004187

Markaki, Yvonni and Simonetta Longhi. 2013. "What Determines Attitudes to Immigration in European Countries? An Analysis at the Regional Level." *Migration Studies* 1(3):311-37. https://doi.org/10.1093/migration/mnt015

Martén, Linna, Jens Hainmueller and Dominik Hangartner. 2019. "Ethnic Networks Can Foster the Economic Integration of Refugees." *Proceedings of the National Academy of Sciences* 116(33):16280-85. https://doi.org/10.1073/pnas.1820345116

Montgomery, Jacob M., Brendan Nyhan and Michelle Torres. 2018. "How Conditioning on Posttreatment Variables Can Ruin Your Experiment and What to Do about It." *American Journal of Political Science* 62(3):760-75. https://doi.org/10.1111/ajps.12357

Muñoz, Jordi, Albert Falcó-Gimeno and Enrique Hernández. 2020. "Unexpected Event during Survey Design: Promise and Pitfalls for Causal Inference." *Political Analysis* 28(2):186-206. https://doi.org/10.1017/pan.2019.27

Newman, Benjamin J. 2013. "Acculturating Contexts and Anglo Opposition to Immigration in the United States." *American Journal of Political Science* 57(2):374-90. https://doi.org/10.1111/j.1540-5907.2012.00632.x

Olzak, Susan. 1994. *The Dynamics of Ethnic Competition and Conflict*: Stanford University Press.

Ondrich, Jan, Stephen L. Ross and John Yinger. 2000. "How Common is Housing Discrimination? Improving on Traditional Measures." *Journal of Urban Economics* 47(3):470-500. https://doi.org/10.1006/juec.1999.2148

Pettigrew, T. F. and L. R. Tropp. 2006. "A meta-analytic test of intergroup contact theory." *Journal of Personality and Social Psychology* 90(5):751-83. https://doi.org/10.1037/0022-3514.90.5.751

Phelps, Edmund S. 1972. "The Statistical Theory of Racism and Sexism." *The American Economic Review* 62(4):659-61.

Phillips, David C. 2019. "Do Comparisons of Fictional Applicants Measure Discrimination When Search Externalities are Present? Evidence from Existing Experiments." *The Economic Journal* 129(621):2240-64. https://doi.org/10.1111/ecoj.12628

Quillian, L., D. Pager, O. Hexel and A. H. Midtbøen. 2017. "Meta-analysis of field experiments shows no change in racial discrimination in hiring over time." *Proceedings of the National Academy of Sciences* 114(41):10870-75. https://doi.org/10.1073/pnas.1706255114

Quillian, Lincoln, Anthony Heath, Devah Pager, Arnfinn H. Midtbøen, Fenella Fleischmann and Ole Hexel. 2019. "Do Some Countries Discriminate More Than Others? Evidence From 97 Field Experiments of Racial Discrimination in Hiring." *Sociological Science* 6:467-96. https://doi.org/10.15195/v6.a18

Rose, Damaris 2019. *Creating a Home in Canada: Refugee Housing Challenges and Potential Policy Solutions*. Washington, DC.: Migration Policy Institute.

Ross, Stephen L. 2017. "Measuring Trends in Discrimination With Field Experiment Data." *Proceedings of the National Academy of Sciences* 114(41):10815. https://doi.org/10.1073/pnas.1715243114

Rudolph, Lukas and Markus Wagner. 2022. "Europe's Migration Crisis: Local Contact and Out-Group Hostility." *European Journal of Political Research* 61:268-80. https://doi.org/10.1111/1475-6765.12455

Scacco, Alexandra and Shana S. Warren. 2018. "Can Social Contact Reduce Prejudice and Discrimination? Evidence from a Field Experiment in Nigeria." *American Political Science Review* 112(3):654-77. https://doi.org/10.1017/s0003055418000151

Schaub, Max, Johanna Gereke and Delia Baldassarri. 2020. "Strangers in Hostile Lands: Exposure to Refugees and Right-Wing Support in Germany's Eastern Regions." *Comparative Political Studies* 54(3–4): 686–717. https://doi.org/10.1177/0010414020957675

Schmid, Laura. 2015. "Ethnische Diskriminierung bei der Wohnungssuche: Feldexperimente in sechs deutschen Großstädten [Dissertation Thesis]." University of Konstanz, Konstanz.

Schmidt, Peter, Stefan Weick and Daniel G. Gloris. 2019. "Wann wirken Kontakte zwischen Migranten und Mehrheitsgesellschaft? Längsschnittanalysen zu Erfahrungen mit Kontakten und zur Bewertung von Flüchtlingen und Muslimen durch die deutsche Bevölkerung." *Informationsdienst Soziale Indikatoren* 61:24-29.

Semyonov, Moshe, Rebeca Raijman, Anat Yom Tov and Peter Schmidt. 2004. "Population Size, Perceived Threat, and Exclusion: A Multiple-Indicators Analysis of Attitudes Toward Foreigners in Germany." *Social Science Research* 33(4):681-701. https://doi.org/10.1016/j.ssresearch.2003.11.003

Shadish, William R., Thomas D. Cook and Donald T. Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, MA: Houghton Mifflin.

Steinmayr, Andreas. 2021. "Contact versus Exposure: Refugee Presence and Voting for the Far Right." *The Review of Economics and Statistics* 103(2):310-27. https://doi.org/10.1162/rest_a_00922

Torres, Katia Gallegos. 2022. "The 2015 Refugee Inflow and Concerns Over Immigration." *European Journal of Political Economy*:102323. https://doi.org/10.1016/j.ejpoleco.2022.102323

UNHCR, The UN Refugee Agency 2022. *Global Trends. Forced Displacement in 2021*. Copenhagen, Denmark: United Nations High Commissioner for Refugees, UNHCR Global Data Service.

Van der Bracht, Koen, Ad Coenen and Bart Van de Putte. 2015. "The Not-In-My-Property Syndrome: The Occurrence of Ethnic Discrimination in the Rental Housing Market in Belgium." *Journal of Ethnic and Migration Studies* 41(1):158-75. https://doi.org/10.1080/1369183x.2014.913476

Vanderweele, Tyler J. 2015. *Explanation in Causal Inference : Methods for Mediation and Interaction*. Oxford: Oxford University Press.

Wagner, Ulrich, Sarantis Tachtsoglou, Patrick Ferdinand Kotzur, Maria-Therese Friehs and Uwe Kemmesies. 2020. "Proportion of Foreigners Negatively Predicts the Prevalence of

Xenophobic Hate Crimes within German Districts." *Social Psychology Quarterly* 83(2):195-205. https://doi.org/10.1177/0190272519887719

Weichselbaumer, Doris. 2015. "Testing for Discrimination against Lesbians of Different Marital Status: A Field Experiment." *Industrial Relations: A Journal of Economy and Society* 54(1):131-61. https://doi.org/10.1111/irel.12079

Zhang, Nan, Johanna Gereke and Delia Baldassarri. 2022. "Everyday Discrimination in Public Spaces: A Field Experiment in the Milan Metro." *European Sociological Review* 38(5):679-93. https://doi.org/10.1093/esr/jcac008

Zolberg, Aristide R. and Long Litt Woon. 1999. "Why Islam is like Spanish: Cultural Incorporation in Europe and the United States." *Politics & Society* 27(1):5-38. https://doi.org/10.1177/0032329299027001002

Zschirnt, Eva. 2019. "Research Ethics in Correspondence Testing: An Update." *Research Ethics* 15(2):1-21. https://doi.org/10.1177/1747016118820497

Katrin Auspurg: LMU Munich, Department of Sociology
E-mail: katrin.auspurg@lmu.de.

Renate Lorenz: LMU Munich, Department of Sociology
E-mail: renate.lorenz@soziologie.uni-muenchen.de.

Andreas Schneck: LMU Munich, Department of Sociology
E-mail: andreas.schneck@lmu.de.