DISCUSSION PAPER SERIES

DP18113

DISCRIMINATION ON THE CHILD CARE MARKET: A NATIONWIDE FIELD EXPERIMENT

Henning Hermes, Philipp Lergetporer, Fabian Mierisch, Frauke Peter and Simon Wiederhold

LABOUR ECONOMICS AND PUBLIC ECONOMICS

ISSN 0265-8003

DISCRIMINATION ON THE CHILD CARE MARKET: A NATIONWIDE FIELD EXPERIMENT

Henning Hermes, Philipp Lergetporer, Fabian Mierisch, Frauke Peter and Simon Wiederhold

Discussion Paper DP18113
Published 24 April 2023
Submitted 18 April 2023

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Labour Economics
- Public Economics

Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Henning Hermes, Philipp Lergetporer, Fabian Mierisch, Frauke Peter and Simon Wiederhold

DISCRIMINATION ON THE CHILD CARE MARKET: A NATIONWIDE FIELD EXPERIMENT

Abstract

We provide the first causal evidence of discrimination against migrants seeking child care. We send emails from fictitious parents to > 18, 000 early child care centers across Germany, asking if there is a slot available and how to apply. Randomly varying names to signal migration background, we find that migrants receive 4.4 percentage points fewer responses. Responses to migrants also contain substantially fewer slot offers, are shorter, and less encouraging. Exploring channels, discrimination against migrants does not differ by the perceived educational background of the email sender. However, it does differ by regional characteristics, being stronger in areas with lower shares of migrants in child care, higher right-wing vote shares, and lower financial resources. Discrimination on the child care market likely perpetuates existing inequalities of opportunities for disadvantaged children

JEL Classification: J13, J18, J22, C93

Keywords: Child care, Discrimination, Information provision, Inequality, Field experiment

Henning Hermes - hermes@dice.hhu.de
DICE (Düsseldorf Institute for Competition Economics), HHU Düsseldorf and CEPR

Philipp Lergetporer - philipp.lergetporer@tum.de
Technical University of Munich, TUM School of Management, Heilbronn & ifo Institute Munich

Fabian Mierisch - Fabian.Mierisch@ku.de
KU Eichstätt-Ingolstadt, Ingolstadt School of Management & ifo Institute Munich

Frauke Peter - peter@dzhw.eu
German Centre for Higher Education Research and Science Studies (DZHW) and DIW Berlin

Simon Wiederhold - simon.wiederhold@iwh-halle.de
Halle Institute for Economic Research (IWH), MLU Halle-Wittenberg and ifo Institute Munich

Acknowledgements
We thank Martin Abel, Mallory Avery, Vojtech Bartoš, Stefan Bauernschuster, Peter Bergman, Aislinn Bohren, Alexander Cappelen, Mathias Ekström, Ruben Enikolopov, Katharina Hartinger, Boon Han Koh, Giovanni Peri, Mirco Tonin, Bertil Tungodden, and seminar participants at the ASSA, VfS, EALE, AFE, SOLE, RES, In_equality Conference Konstanz, Applied Young Economists Webinar, CESifo Area Confernce, DICE Düsseldorf, FAIR NHH Bergen, KU Leuven, ifo Institute Munich, and KU Eichstaett-Ingolstadt for valuable comments and suggestions. Lukas Seubert provided excellent research assistance. We also thank David Broneske and Nikhilkumar Italiya for their support in implementing the natural language processing model. We gratefully acknowledge funding from the Technical University of Munich, the DICE Düsseldorf, the KU Eichstätt-Ingolstadt (ProFOR+ funding

No. U060230003FL 4/3), and the ifo Institute Munich. The study was preregistered in the AEA RCT Registry (AEARCTR-0007389). IRB approval was obtained from the Ethics Commission of the Department of Economics, University of Munich (Project 2020-17).

Discrimination on the Child Care Market: A Nationwide Field Experiment*

Henning Hermes[a], Philipp Lergetporer[b], Fabian Mierisch[c], Frauke Peter[d], Simon Wiederhold[e]

This version: April 3, 2023

[a] *DICE (Düsseldorf Institute for Competition Economics), HHU Düsseldorf*
[b] *Technical University of Munich, TUM School of Management, Heilbronn & ifo Institute Munich*
[c] *KU Eichstätt-Ingolstadt, Ingolstadt School of Management*
[d] *German Centre for Higher Education Research and Science Studies (DZHW) & DIW Berlin*
[e] *Halle Institute for Economic Research (IWH), MLU Halle-Wittenberg & ifo Institute Munich*

Abstract

We provide the first causal evidence of discrimination against migrants seeking child care. We send emails from fictitious parents to $> 18,000$ early child care centers across Germany, asking if there is a slot available and how to apply. Randomly varying names to signal migration background, we find that migrants receive 4.4 percentage points fewer responses. Responses to migrants also contain substantially fewer slot offers, are shorter, and less encouraging. Exploring channels, discrimination against migrants does not differ by the perceived educational background of the email sender. However, it does differ by regional characteristics, being stronger in areas with lower shares of migrants in child care, higher right-wing vote shares, and lower financial resources. Discrimination on the child care market likely perpetuates existing inequalities of opportunities for disadvantaged children.

Keywords: child care, discrimination, information provision, inequality, field experiment

JEL: J13, J18, J22, C93

*We thank Martin Abel, Mallory Avery, Vojtěch Bartoš, Stefan Bauernschuster, Peter Bergman, Aislinn Bohren, Alexander Cappelen, Mathias Ekström, Ruben Enikolopov, Katharina Hartinger, Boon Han Koh, Giovanni Peri, Mirco Tonin, Bertil Tungodden, and seminar participants at the ASSA, VfS, EALE, AFE, SOLE, RES, In_equality Conference Konstanz, Applied Young Economists Webinar, CESifo Area Confernce, DICE Düsseldorf, FAIR NHH Bergen, KU Leuven, ifo Institute Munich, and KU Eichstaett-Ingolstadt for valuable comments and suggestions. Lukas Seubert provided excellent research assistance. We also thank David Broneske and Nikhilkumar Italiya for their support in implementing the natural language processing model. We gratefully acknowledge funding from the Technical University of Munich, the DICE Düsseldorf, the KU Eichstätt-Ingolstadt (ProFOR+ funding No. U060230003FL 4/3), and the ifo Institute Munich. The study was preregistered in the AEA RCT Registry (AEARCTR-0007389). IRB approval was obtained from the Ethics Commission of the Department of Economics, University of Munich (Project 2020-17).

1. Introduction

Early child care is a key policy tool for promoting children's development and long-term economic success (e.g., Currie and Almond, 2011; Heckman et al., 2013). Especially children with migration background benefit from early child care, for example, in terms of language skills and school readiness (Cornelissen et al., 2018; Felfe and Lalive, 2018). Therefore, early child care has the potential to foster social equality by leveling the playing field. However, migrant children are underrepresented in early child care in many countries: The average enrollment gap between children with migrant and native parents across OECD countries is 12 percentage points (pp), equivalent to 36% of the native enrollment rate (OECD, 2018). These disparities prompt the fundamental question of *why* children with migration background are less frequently enrolled in early child care.

Demand-side factors are unlikely to be solely responsible for the migrant-native gap in early child care enrollment. In Germany, the country we study, demand for an early child care slot is the same for migrant and native parents, but migrant children are 12 pp less likely to be enrolled (Jessen et al., 2020). Similarly, when comparing only those parents who actually have applied for child care, migrant parents have a lower likelihood of securing a slot compared to native parents (Hermes et al., 2021). It is also unlikely that parents' financial constraints explain the enrollment gap, as this gap also exists in countries with minimal or no child care fees (see OECD, 2018, 2020).

Thus, there is reason to suspect that *supply-side factors* matter for explaining the migrant-native enrollment gap in early child care. Similar to many other developed countries (see Section 2), child care in Germany is universal, i.e., targeted to all children, with high quality and low fees. Nonetheless, there is competition for available slots, as parental demand exceeds the actual supply of child care slots. Additionally, the application process for early child care is complex and non-transparent, which makes it imperative for parents to acquire information about when, where, and how to apply to secure a slot. Such information is often provided by the local child care center managers. At the same time, these local center managers are also in charge of the admission decisions, with no binding admission criteria or accountability systems in place. This decentralized admission system gives center managers substantial discretion in deciding whom to provide information and whom to enroll. However, despite this key role of center managers in the child care application process, it has not yet been studied whether their behavior contributes to the migrant-native enrollment gap.

We conduct a nationwide field experiment with more than half of the early child care centers ($> 18,000$) in Germany to examine whether center managers discriminate against families with a migration background.[1] We send emails from fictitious parents drafted according to a set of real email messages from parents. The emails ask if there are slots available and how to apply. These are the two most frequent questions parents ask when contacting child care centers, as a preceding survey with more than 400 child care center managers showed (see Appendix C.1). In the emails, we randomly vary whether or not a parent's name signals a migration background (Bertrand and Mullainathan, 2004). We chose the most common German (*native*) and Turkish (*migrant*) names based on German name registry data, and validated our name selection in additional pretests (see Appendix C.2). Turkish migrants are particularly relevant in the German context, as they are the largest and geographically most widespread migrant group. Despite often living in Germany for generations, Turkish families are strongly underrepresented in early child care, while their demand is very similar compared to natives (see Appendix D for details).

We find a large negative effect of a migrant sender name on the response rate. Emails signaling that the sender is a migrant receive 4.4 pp fewer replies from child care center managers than emails signaling that the sender is a native. With a response rate of 71% for native senders, this translates to a scaled effect of -6.2%. This effect size is large compared to other comparable correspondence studies. For instance, Hemker and Rink (2017) found no migrant-native gap in response rates in their correspondence study with German welfare offices. In recent large U.S. correspondence studies, estimated black-white response gaps are around 2 to 4 pp (see, e.g., Bergman and McFarlin, 2018; Giulietti et al., 2019; Kline et al., 2022). Hence, our results demonstrate that migrants in Germany face substantial discrimination when searching and applying for child care.

Because we receive written responses, we can analyze not only response rates, but also the *content* of the email responses. The most important dimension of the email content is whether parents are actually offered a slot in the child care center. We find that email replies to a migrant sender are significantly less likely to contain actual slot offers than replies to a native sender. Unconditional on receiving a response, the treatment effect amounts to -1.1 pp, which corresponds to -23% relative to the mean offer rate for natives. Conditional on receiving a response, the effect is only slightly smaller at -1.0 pp or -20.5%. While these offers are non-binding and mostly require further action by the parents, this

[1] Our definition of discrimination follows Bertrand and Duflo (2017), p. 310: "Members of a minority group [...] are treated differentially (less favorably) than members of a majority group with otherwise identical characteristics in similar circumstances." Applied to our setting, this means that migrant parents have a lower likelihood of receiving a response to the exact same email, or receive a response with less positive content, than native parents.

result suggests that discrimination against migrants on the early child care market is not limited to information provision, but also extends to the allocation of slots.

Consistently, we also find migrant-native gaps on other email content dimensions: Both conditional and unconditional on receiving a response, email replies to a migrant sender are shorter, contain fewer offers to join a waiting list, and are rated as less helpful, less encouraging, and generally less appealing in substance and tone. Overall, our results show that discrimination against migrants exists at both the extensive margin (response rates) and the intensive margin (content of the email).

Next, we investigate potential mechanisms underlying discrimination of child care center managers against migrant families. First, we causally examine whether discrimination is based on child care center managers' beliefs about parents' educational background. To do so, we independently randomize whether the email contains a signature indicating that the sender has a university degree. Adding this higher education signal does not significantly affect the migrant-native response rate gap. We therefore consider it unlikely that the lower response rates to migrant parents are due to child care center managers discriminating against a (perceived) lower educational background.

We also explore additional mechanisms by leveraging rich administrative data on the regions where child care centers are located. We investigate three channels: First, in line with the contact hypothesis (Allport, 1954), we find that discrimination is weaker in areas with a larger share of migrant children enrolled in child care. Second, discrimination is stronger in regions with higher right-wing vote shares, showing that general attitudes towards migrants (by center managers or parents of other children) are linked to discrimination on the child care market. Third, child care centers with more resources and those located in federal states that provide financial incentives to enroll migrant children discriminate less. Thus, center managers may perceive migrant children as costlier to educate. Reassuringly, the pattern of these treatment effect heterogeneities is very similar when considering response rates or actual slot offers.

We conduct a variety of tests to confirm the robustness of our main findings. Treatment effects are robust to using (i) corrections for multiple hypothesis testing, (ii) randomization inference, (iii) Probit estimations, and (iv) models with fixed effects for zip codes of the child care centers. Furthermore, while our baseline results for the email content are based on a manual coding of the email replies by a team of five research assistants ("reviewers") blind to the treatment, employing an alternative approach to classify email content based on a supervised machine-learning technique yields similar results.

Our paper contributes to the existing literature in several ways. We add to the literature on discrimination in the education system (e.g., Alesina et al., 2018; Carlana, 2019; Alan et al., 2021; Lavy et al., 2022). While previous studies have documented discrimination in the admission processes of schools (Bergman and McFarlin, 2018; de Lafuente, 2021) and universities (Dynarski et al., 2018; Arcidiacono et al., 2022), we are the first to show the existence of discrimination already very early in the education system, i.e., in the admission to early child care. Doing so, we provide a possible reason for the underrepresentation of migrant children in child care systems worldwide (OECD, 2018; Hussar et al., 2020; Jessen et al., 2020). This underrepresentation is a major concern as it can put migrant children on a worse educational trajectory with persistent negative impacts over their life course (e.g., Heckman et al., 2010).

Furthermore, we document that discrimination occurs not only in official admission procedures but already in the initial stages of gathering information about the application process (see Bartoš et al., 2016; Bergman and McFarlin, 2018). Emails serve as a cheap, easy, and informal way to request information from any type of institution (public or private), and our study reveals that discrimination can manifest even in these early interactions. As emails often serve as the first point of contact between parents and child care centers, such discrimination can have significant consequences by discouraging applications or enrollment. This is especially true in settings where slots are rationed, enrollment procedures are complex, and relevant information is difficult to access — which is characteristic of many child care systems worldwide (Mocan, 2007; Spiess, 2008; Hermes et al., 2021). Importantly, this form of discrimination is very difficult to detect, as there are virtually no recorded data on such informal information inquiries (in contrast to data on applications or enrollment).

Finally, we also add to the extensive literature that uses correspondence studies for detecting discrimination (for reviews, see Baert, 2018; Neumark, 2018). As we can analyze actual email replies, we are able to document that discrimination not only exists for response rates, but also for various email content measures (including objective measures such as slot offers and subjective measures such as the perceived helpfulness of the email). Intriguingly, we show that discrimination is substantial not only at the extensive margin (response rates), but also at the intensive margin (email content), suggesting that relying only on call-back rates (i.e., the main focus of standard correspondence studies) underestimates the true extent of discrimination. The fact that we find discrimination with respect to actual slot offers is also reassuring in light of the criticism against correspondence studies that observed outcomes (e.g., callbacks to job applications) do not entail direct economic consequences, and are merely a proxy for the outcome of interest (e.g., actual job offers, see Bertrand and Duflo, 2017).

The remainder of this paper is structured as follows. Section 2 provides information on the institutional background of the early child care market in Germany. Section 3 describes our experimental design and the data. Section 4 presents our empirical strategy. Section 5 reports our results, the mechanism analysis, and robustness checks. Section 6 concludes.

2. Institutional Background

In this section, we discuss several key features of the German child care system that are potentially conducive to discrimination. As shown in Table A1, these institutional features are similar in many other developed countries.

Early child care provision. Child care provision in Germany is universal, targeting all children before they enter school (at the age of six). Child care is available for children at two distinct age groups: (i) under the age of three years (early child care or *Krippe*) and (ii) between three and six years (*Kindergarten*). Each child has a legal entitlement to a child care slot from the age of one year onward. Early child care is heavily publicly subsidized, with the public sector paying about three-quarters of the total cost (Spiess, 2013). Parents pay very low child care fees (on average 250 EUR per month, equivalent to 10% of the average income), and lower-income families are eligible for fee reductions or even exemptions (Felfe and Lalive, 2018). Compared to other countries, the quality of early child care is relatively high and homogeneous across Germany, for example, in terms of group sizes or staff-to-child ratios (Felfe and Lalive, 2018).

Despite the fact that child care is coined universal, its utilization is far from universal. On average, 34% of children under the age of three are enrolled in early child care. Attendance rates increase substantially with age, from only 1% for children under the age of one to 55% for children aged two to three years (Bildungsberichterstattung, 2020). More than 90% of children attend *Kindergarten*, so that almost all children have attended child care by the time they start school. As a consequence, the relevant margin is not *whether* children have access to child care, but rather *when* they take up child care, in particular, whether they enroll into early child care. Past research has shown that earlier enrollment in child care can have pronounced positive effects on child development (Drange and Havnes, 2019).

Like in many other countries, a key characteristic of the German child care market is the rationing of slots, as parental demand exceeds supply. Rationing is especially prevalent for parents with a migration background. While there is almost no difference between native and migrant parents in the wish to enroll their child in early child care (Jessen et al., 2020), actual enrollment differs substantially: Only 24% of children with a migration background are enrolled, compared to 38% of native children (Schmitz et al., 2023). Note that from a purely legal perspective, this migrant-native enrollment gap should not be due to discrimination against migrants, because it is illegal in Germany to deny children access to child care due to their migration background, as the General Equal Treatment Act prohibits discrimination on the basis of race, ethnic origin, or other characteristics by constitutional law (*Allgemeines Gleichbehandlungsgesetz*). However, whether discrimination occurs in reality is an empirical question, which we address in this paper.

Organization and funding of early child care. In Germany, child care is part of the child and youth welfare system under the responsibility of the federal government but the actual implementation of child care provision takes place at the municipality level. As such, these decentralized, local child care markets are vastly heterogeneous and differ in prices for child care, fee reductions, application procedures and deadlines, and admission criteria.[2] Child care centers are mostly government-funded, and thus depend only little on fees paid by parents (Alt et al., 2019). Across regions, there are substantial differences in the amount of resources and the staff-to-child ratios in early child care centers. Moreover, in nine out of 16 federal states, centers receive additional funding to care for children with a migration background. These additional funds are mainly intended to promote children's language skills.

In Germany, 83% of child care centers are directly operated by municipalities or non-profit organizations (ecclesiastical or charitable organizations), 15% are operated by different associations (parents, youth, and others), and 3% are run by for-profit organizations or companies (see Bildungsberichterstattung, 2020). Child care is mostly provided by small centers (typically serving to 25–75 children, see DJI, 2021), and competition between centers is low (Spiess, 2008). Importantly, in almost all German child care centers, managerial tasks — such as communicating with parents and allocating slots — are performed by single center manager (79%) or a team of manager and deputy manager (13%), who are released from child care responsibilities to perform these tasks (DJI, 2021).

[2] As a consequence, there is no national accountability system in place, such as the National Association for the Education of Young Children in the United States, on which parents can base their assessment and selection of child care centers.

Enrollment process in early child care. Due to the decentralized structure of the child care market, each center typically has its own enrollment process, resulting in an unstructured and individualized application process for families. Additionally, there are no mandatory, standardized criteria for child care center managers how to prioritize when allocating slots, and no accountability system to track enrollment decisions.[3] Therefore, the application process for child care in Germany is complex and can differ for each child care center, which implies that information about the application process is crucial for parents to successfully enroll their child. Acquiring a slot is likely more difficult for parents with a migration background, because the process of searching and applying for child care is resource-intensive (e.g., in terms of social capital and time). Therefore, obtaining information about how, when, and where to apply for child care is particularly important for migrant parents, who are less likely to have such information compared to native parents (Hermes et al., 2021).

For these reasons, we expect that child care centers (i.e., the supply side of the child care market) are relevant for explaining the large migrant-native gap in early child care enrollment. The absence of a centralized accountability system and transparent admission criteria, as well as the decentralized decision-making at the child care center level, provide center managers with a great amount of discretion in allocating slots.

3. Experimental Design and Data

To investigate discrimination in the German child care market, we conduct an email correspondence study (for recent overviews, see Baert, 2018; Neumark, 2018).[4] We send emails over a three-day period in March 2021, and collect responses for a period of more than two months. In our emails, we experimentally vary (i) the names of the fictitious parents (Bertrand and Mullainathan, 2004; Bertrand and Duflo, 2017) and (ii) whether or not the email contains a signature with an higher education signal (Giulietti et al.,

[3]In our survey, 86% of child care center managers report that the center manager is responsible for enrollment decisions, while less than 10% indicate that admission criteria of the provider or the municipality play a role for these decisions (see Appendix C.1).

[4]Because it is not possible to obtain informed consent from study participants, the barriers for ethical approval of correspondence studies are high. We received IRB approval and clearance in terms of compliance with German data protection rules (DSGVO) from the ethics committee of LMU Munich (reference 2020-17).

2019). We then record whether we receive any response to our email and also analyze the content of the responses received.[5]

3.1. Email Design

To inform our study design, we first conducted qualitative interviews and a nationwide online survey ($n = 447$) with child care center managers (see Appendix C.1 for details). The data reveal that most parents get in touch with child care centers prior to submitting an application, with email being one of the most common communication channels, especially in the early stages of contact. In fact, 85% of child care centers in the survey frequently receive emails from parents (on average, 3.5 emails per week). Requests for open slots (89%) and how to apply (60%) are the two most common questions center managers receive. Finally, about four out of ten (38%) center managers state that email contact with parents is an important prerequisite for actual enrollment, highlighting the relevance of email requests for the allocation of (rationed) slots in child care.

We drafted the email based on (i) the information from our survey with center managers and (ii) a set of 12 real, anonymized email messages from parents which we obtained confidentially from a child care center. We identified similarities in these emails in terms of sentence structure, length, and content for our email. Based on the center manager survey, we selected the two questions that are most frequently asked by parents: whether there is an open slot and how to apply. Next, we recruited a sample of online workers ($n = 200$) to rate the degree of realism of our fictitious email.[6] Reassuringly, 80% of survey participants rated the email as realistic or very realistic, whereas only 4% rated the email as not realistic. The high overall response rate in our study indicates that our efforts to design a realistic email were successful.

Figure 1 presents the general email template; specific examples are shown in Appendix Figures A1 and A2. To ensure that our findings are not influenced by any gender-related response biases from child care center managers, we randomly varied the gender of both the child and parent in the emails we sent.[7] The email indicates that the child is one year and five months old. The parent then asks (i) if the child care center has a slot for the

[5] While other correspondence studies often contact study participants multiple times, we sent only one email to each child care center to impose as little cost as possible on them and to minimize detection risk (Bertrand and Duflo, 2017).

[6] The survey was conducted using Clickworker, a German online platform similar to Amazon MTurk (see Appendix C.2 for further information).

[7] In Hermes et al. (2023), we use the sample of emails with a native sender to analyze how the parent's gender affects response behavior.

Figure 1: Email Template with Randomized Information Highlighted

> Dear Sir or Madam,
>
> We are looking for a child care slot for our [son/daughter] starting in January 2022. [He/she] is now 1 year and 5 months old.
>
> Do you still have a slot available? How can we apply for a slot?
>
> Thank you!
>
> Sincerely,
> [Name]
>
> ---
> [Name], Bachelor of Arts (FH)
> Email: [Name]@ ...
> ---

Notes: Figure shows the email sent from fictitious parents to child care centers, translated from the original German version. Text marked in grey is randomized and differs by version of the email. We randomized the gender of the child (2 variations), the name of the parent (16), and whether or not we include an email signature (2). The signature indicates that the parent holds a bachelor's degree from a University of Applied Sciences, which is the most common university degree in Germany. We sent a total of 64 ($2 \times 16 \times 2$) different versions of the email to child care centers.

child in about nine months (i.e., when the child will be over two years old; some centers do not accept children before this age cut-off), and (ii) how to apply for a slot.

Note that we intentionally designed our email to be free of grammatical errors, typos, or other mistakes that might (statistically) be correlated with sender characteristics. Since we expect such formal errors to be more likely to occur among migrants (e.g., due to a lower language proficiency), any treatment effects should be interpreted as a lower bound estimate for the discrimination experienced by actual migrants.[8]

3.2. Treatment Variations

To signal whether the parents seeking child care had a native or migration background, we experimentally varied the name of the email sender (Bertrand and Mullainathan, 2004). We selected names that most clearly signal a native or Turkish migration background in

[8]Note that in our survey of online workers, we do not find significant differences in realism ratings of emails by migration background (or by gender).

the following way: First, we chose the most common names for native Germans living in Germany and for Turkish migrants living in Germany, respectively.[9] Second, we used our survey of online workers to elicit which country of origin respondents associated with different first names and surnames. Finally, among the names for which at least 90% of respondents indicated a German or Turkish background, we selected two first names for each gender and two surnames for "German" and "'Turkish" senders (see Table 1).[10]

In total, we sent emails from 16 fictitious parents (four female first names, four male first names, each with two different surnames). We created an email account for each of the 16 fictitious parent names (e.g., Andreas Schmidt, Fatma Öztürk) following the pattern of name.surname0528@.... Each account sent about 1,400 emails (see Section 3.3).[11]

We chose Turkish names to indicate migration background because Turkish immigrants are by far the largest and geographically most dispersed migrant group in Germany (around 13% of all migrants are Turkish; for additional information, see Appendix D). People with a Turkish migration background have often lived in Germany for several decades, as the major emigration waves from Turkey to Germany took place in the 1970s. At the same time, children from Turkish families are strongly underrepresented in early child care. In fact, the enrollment rate for Turkish children in early child care is 12%, compared to 33% for German children (see Jessen et al., 2020).

Our second treatment variation aimed to fix child care center managers' beliefs about the educational background of the sender.[12] For instance, if Turkish names were associated with lower levels of education, center managers might be more likely to discriminate

[9]We used the online portal of the Society for German Language (GfdS) to identify the most common names from the cohort born in 1986.

[10]Using data from a public website providing ratings for first names, we can also rule out that any of the names selected for our study are outliers in terms of their popularity ratings.

[11]In one of the 16 accounts, there was a spelling mistake: while the *email address* was correctly shown as oemer.yildirim0528@..., the *sender name* and, depending on the treatment, the email signature displayed "Yildrim" instead of "Yildirim" as surname. The response rate to this account was lower than for the other accounts (about 20 pp on average), and we cannot rule out the possibility that this was due to the spelling mistake. In our regressions, we thus always include a dummy for this email account and an interaction term with the higher education signal. Naturally, our treatment effect estimates become even stronger if we do not control for the account. Results are virtually identical if we exclude emails sent from this account from the estimation sample.

[12]We decided to vary parents' educational background instead of other potential mechanisms for three main reasons: First, education is a commonly used proxy for socioeconomic background more generally (see Bjoerklund and Salvanes, 2011), and thus covers a broad range of relevant parental characteristics. Second, educational background is one of the best predictors of parental engagement in the child care center — a margin that is likely to be critical for center managers' decisions about slot allocation. Third and pragmatically, educational background can easily be signaled in emails in an unobtrusive way (in contrast to, say, occupations, which are very rarely mentioned in actual emails).

Table 1: Names Used to Signal Native and Turkish Migration Background

	German	**Turkish**
Male	Andreas Sebastian	Ömer Hüseyin
Female	Stefanie Christina	Eylül Fatma
Surname	Schmidt Müller	Yildirim Öztürk

against such senders simply because of the (perceived) lower educational background. To test this possibility, a random subset of the sent emails included an email signature indicating that the sender has obtained a university degree (Giulietti et al., 2019). For the signature, we used the most frequently obtained university degree in the German population, a Bachelor of Arts from a University of Applied Sciences (HRK, 2021). Since approximately 30% of individuals in the relevant age group (30–34 years) in Germany hold this university degree, our treatment effectively conveys an above-average level of education while still representing a sizeable portion of the population. Using a higher degree, such as a Ph.D., which is held by only 2% of the population, would not have been as representative. Through our survey of online workers, we verified that respondents were able to recall the information from the email signature indicating that the sender possessed a university degree (see Appendix C.2).

We varied the migrant treatment and the higher education signal independently in a 2×2 design, resulting in four different treatment cells. In total, we sent 64 emails combining the 16 names, the higher education signal, and the gender of the child.

3.3. Sample

For the sampling, we use a comprehensive and commercially available data set of child care centers with email addresses, covering the vast majority of all child care centers in Germany. Because we focus on early child care, we restrict the full sample to centers that enroll children below the age of three years. We further exclude about 2,000 centers that share an email address with another center to minimize the risk of detection. We end up with a sample of 22,458 child care centers for our field experiment, representing about 60% of all early child care centers in Germany (Destatis, 2020).

We sent a total of $n = 22,458$ emails to an equal number of distinct early child care centers. Of these emails, $n = 3,795$ were not delivered to the recipient (mostly because the email address was no longer valid).[13] Our final analysis sample consists of $n = 18,663$ delivered emails. We received a total of $n = 12,547$ responses, resulting in an average response rate of 67.2%.[14]

We use stratified randomization to generate the four treatment groups. We construct strata by federal state, level of urbanization of the county where the center is located (i.e., predominantly urban, intermediate, or predominantly rural), and provider type (i.e., public, church-run, and other types of child care centers, such as centers run by for-profit providers or by parental initiatives). In total, we use 131 strata in our randomization.

Table A2 shows summary statistics and balancing in the analysis sample (see Appendix E for definitions of all variables used). Half of the child care centers received emails from male senders, and half received emails indicating a male child. The average child care center has a maximum capacity of 68 children. All centers in our sample provide early child care for children under the age of 3 years, most (93%) also cater for children between 3 and 6 years, and only a few (9%) offer afternoon daycare for children above 6 years. One-quarter of child care centers are operated by church, and one-fifth are public child care centers.

Regarding regional characteristics, 44% of child care center are located in a predominantly urban county, 37% in an intermediate county, and 18% in a predominantly rural county. For our mechanism analysis in Section 5.3, we additionally use regional information on the share of migrant children in child care, right-wing vote shares, staff-to-child ratios, and whether the federal state provides additional funding to care for migrant children. The average migrant share in a county is 23%, while the average share of migrants in child care is somewhat higher at 30%. The staff-to-child ratio is about 1:8, the right-wing vote share is 10%, and 87% of centers are located in a federal state that offers financial incentives from the federal state to enroll migrant children.

Table A2 also shows that our randomization was successful. Despite the large sample size, we cannot detect any statistically significant differences between the characteristics of the baseline group (native sender, no higher education signal) and those of the other three experimental groups—a total of 96 comparisons. All differences are also economically

[13]The likelihood of an email not being delivered is independent of treatment status (see Table A2).

[14]In addition, we received around 400 emails (~2%) which we could not link to an observation in our database. These cases are mostly emails that have been forwarded to youth welfare offices or other child care centers, which then responded to our initial request, and therefore could not be matched to initial emails.

negligible in size. The table further indicates that emails that could not be delivered due to wrong or outdated email addresses are randomly distributed across groups. In Table A3, we repeat the same analysis in the sample of all sent emails (including bounced emails), again showing that the stratified randomization has been successful.

4. Empirical Strategy

4.1. Estimation

Our preregistered main specification estimates treatment effects by regressing the outcome of interest (see Section 4.2) on randomized treatment indicators using ordinary least squares (OLS) models. In order to increase precision and to account for slight imbalances between treatment and control groups, we include a vector of preregistered control variables in our main specification:

$$Y_{ij} = \beta_0 + \beta_1 Migrant_j + \mu Controls_{ij} + \varepsilon_{ij} \tag{1}$$

Here, Y_{ij} is the outcome of interest for a fictitious parent j sending an email to the child care center i. $Migrant_j$ is our main treatment variable, taking a value of one if the name of the parent j signals a Turkish migration background, and zero if it signals a native background. $Controls_{ij}$ is a vector of control variables that includes additional randomly assigned email characteristics (i.e., child and sender gender), child care center characteristics (i.e., center's maximum capacity and indicators of whether the child care center also has a kindergarten or a daycare), characteristics of the municipality in which the contacted child care center is located (i.e., share of migrants in the municipality), and strata fixed effects (i.e., interactions between provider type, urban class, and federal state).

In Appendix E, we provide further details on variable definitions and data sources.[15] ε_{ij} is an idiosyncratic error term. Based on the randomized research design, the causal effect of the migrant treatment on outcome Y_{ij} is given by β_1.

[15] As detailed in Appendix E, there are missing observations for a child care center's maximum capacity (6.4%), its provider type (29.4%), and the migrant share of the center's municipality due to missing information about the center's location (7.0%). We impute missing values for maximum capacity and migrant share by the sample mean of the next-higher higher regional level and add imputation indicators to our regressions. We further assign missing values for the provider type to the "else" category and control for a missing indicator.

To measure treatment effect differences for emails with vs. without the higher education signal, we also estimate the following preregistered regression model:

$$\begin{aligned} Y_{ij} = \; & \gamma_0 \; + \; \gamma_1 Migrant_j \; + \; \gamma_2 HigherEducation_j \\ & + \; \gamma_3 Migrant_j \times HigherEducation_j \; + \; \nu Controls_{ij} \; + \; \upsilon_{ij} \end{aligned} \tag{2}$$

Here, $HigherEducation_j$ is an indicator variable equal to one if the email contains a signature indicating that the sender has obtained a university degree, and zero if no such education signal is included. We add the same vector of controls as in equation (1). In this specification, γ_1 shows the migrant treatment effect with no higher education signal, while γ_2 indicates the effect of the higher education signal for native senders. γ_3 shows whether the migrant treatment effect is different with and without the higher education signal. The migrant treatment effect with the higher education signal is given by $\gamma_1 + \gamma_3$.

In our robustness section, we show that our results hold for alternative specifications and multiple hypothesis testing corrections (see Section 5.4).

4.2. Outcome Measures

The main outcomes of our study are (i) a binary indicator for whether or not a contacted child care center replies to the email request, and (ii) several dimensions of the content of the email response (*Slot Offer*, *Waiting List*, *Long Response*, *Helpful*, *Encouraging*, and *Recommend*; for details see below).

We calculate the length of responses by using an automated method that counts the number of characters in the email body. To measure the other content dimensions, we employed a team of five research assistants (“reviewers”), each of whom rated all 12,547 email responses independently. The reviewers received a half-day training session and a manual that explained how to review each email and code responses. Importantly, reviewers were not informed about the purpose of the study or that there were randomized treatments. Furthermore, reviewers were blind to both the treatment assignment, i.e., the name of the sending parent, as well as the identity of the child care centers or managers, as we deleted all this information from the emails prior to the rating.

For each outcome dimension, we aggregate the five reviewer ratings into a binary indicator using simple majority rule (see Appendix F for details on the rating process and on the variable definitions).[16] To verify the robustness of our email content analysis, we

[16] All results are robust to using alternative aggregation rules (see Section 5.4).

show that results hold when using a supervised machine-learning algorithm to classify the content of email responses instead of human coders (see Section 5.4 and Appendix G).

To capture the various possible dimensions of discrimination that may arise in email responses, we defined the following set of content dimensions which were then rated by each reviewer:

Slot Offer. Our most important content dimension is whether the email response contains a slot offer to parents, as it allows us to uncover discrimination in actual enrollment decisions. Although the email requests only inquired about the availability of slots and the application process, 798 center managers (4.9%) directly offered slots in their response email.[17]

Waiting List. The next content dimension is whether child care centers offer placing families on a waiting list. A place on a waiting list provides the possibility of future enrollment if a slot at the center becomes available. In total, 9,976 email responses (56.6%) include a waiting list offer. Although being placed on a waiting list is a step closer to a slot offer, waiting list offers are non-binding and will often not lead to actual enrollment.

Long Response. The length of the email response serves as a proxy for its informational value. Receiving information about the child care application process is particularly important for migrant parents, as they often lack such information: For instance, in a sample of >600 parents with young children in Hermes et al. (2021), only 61% of migrant parents know that they have a legal entitlement to child care once the child is one year old, compared to 85% of native parents. Similarly, only 77% of migrant parents (compared to 94% of native parents) are aware that low-income families are eligible for child care fee reductions or waivers. We measure response length as the number of characters in the text body (excluding names, signatures, and email histories), and create a dummy indicating above-median response length.

Helpful, Encouraging, Recommend. The final three email content dimensions relate to the tone of the email response, because a negative tone may dissuade parents from applying for a slot at the responding center.[18] It is difficult to measure the tone of unstructured

[17] Note that we declined all slot offers within 24 hours.

[18] In fact, a non-negligible share of parents hold the belief that they can only apply to a single child care center. In the above-mentioned study by Hermes et al. (2021), 43% of parents (erroneously) believe that they are legally obligated to choose the nearest child care center for their children.

textual emails using objective criteria, which is why we rely on subjective assessments from the reviewers. Specifically, we asked them to assess whether the email response was (i) helpful or (ii) encouraging, and, to judge the overall impression of the email, (iii) whether they would recommend a befriended couple seeking child care for their 1.5 year old child to apply to the responding child care center.

An important question in analyzing treatment effects on content measures is whether outcomes should be considered unconditional on whether a child care center responds (i.e., non-response would be coded as zero), or conditional on response (i.e., non-response would be coded as missing). The advantage of unconditional content measures is that estimates are not biased by selection into response. However, the disadvantage is that any treatment effects observed on unconditional content measures may be attributable solely to the impact of the treatment on response rates. Therefore, our content analysis in Section 5.2 presents both, treatment effects unconditional and conditional on response.

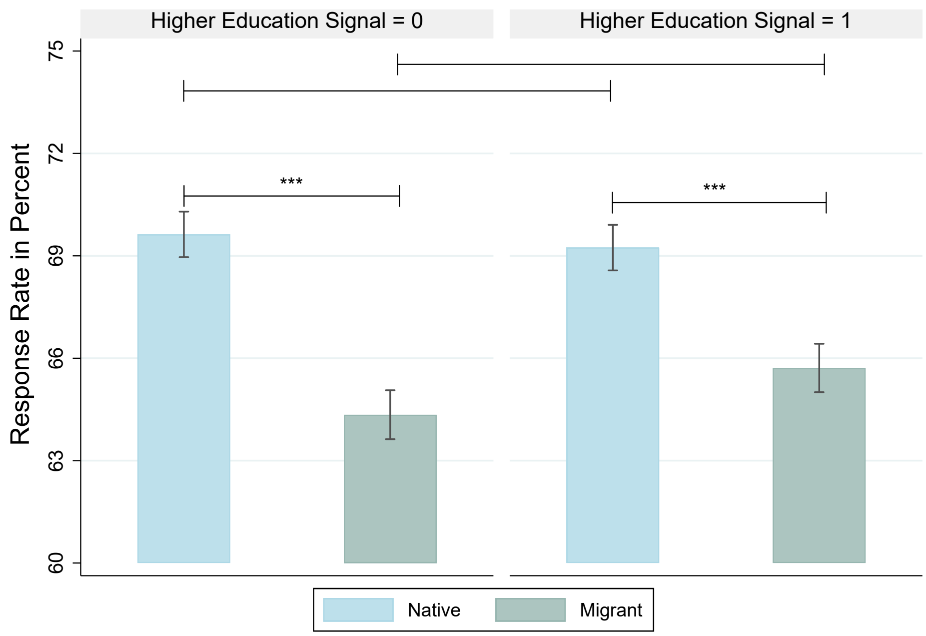
5. Results

5.1. Results for Response Rate

We present our main findings for response rate in Figure 2. The left panel shows response rates to emails sent by parents with a native or a migrant name, without the higher education signal. The mean response rate in the native control group is 69.7% (blue bar). Emails signaling migrant background received 5.3 pp fewer answers (green bar). This difference is highly statistically significant ($p < .0001$). The right panel of Figure 2 shows response rates to emails which include the higher education signal. Comparing the left and right panels reveals two key findings. First, response rates to emails with or without the higher education signal are not significantly different for both natives and migrants ($p = .644$ and $p = .197$, respectively). Second, even when we fix the education level of the senders, the response rate to emails from migrants is 3.5 pp lower than to emails from natives ($p = .0003$). The difference between the migrant treatment effects across both panels is not statistically significant ($p = .203$). Therefore, our results suggest that child care managers' lower response rates towards migrant parents cannot be attributed to their beliefs about the educational background of migrants and natives.

In Table 2, we investigate treatment effects on response rates in an OLS regression framework. In Column (1) and (2), we pool observations with and without the higher education signal (see Equation (1)), while Columns (3) and (4) consider education signal effects (see Equation (2)). The odd columns show raw treatment effects without control

Figure 2: Treatment Effects on Response Rate



Notes: Figure shows response rates across treatment cells, based on multivariate OLS regressions shown in Table 2, Column (4). The left (right) panel depicts response rates to native and migrants senders without (with) a higher education signal. In both panels, the response rate difference between native and migrant senders is statistically significant at the 1%-level. The migrant treatment effects in both panels are not statistically significantly different from each other ($p = .203$). Response rates for native senders with vs. without the higher education signal (the blue bars) and response rates for migrant senders with vs. without the higher education signal (the green bars) are not statistically significantly different from each other ($p = .644$ for natives and $p = .197$ for migrants). Differences are tested with two-sided t-tests. Error bars indicate standard errors. Significance levels: * $p < .10$, ** $p < .05$, *** $p < .01$.

variables, the even columns include the set of control variables. In the pooled specifications in Columns (1) and (2), the migrant treatment effect is -4.4 pp, which translates to a 6.2% decrease relative to the response rate in the native control group. In Column (4), we present the regression analysis that underlies Figure 2, showing migrant treatment effects of -5.3 pp without the higher education signal and -3.5 pp with the higher education signal. The migrant treatment effects do not differ significantly between emails with and without the higher education signal, as indicated by the non-significant interaction term ($p = .203$). Thus, we use the pooled model in Column (2) in subsequent analyses.

In sum, our findings show that migrant parents searching for a child care slot face substantial discrimination at the "extensive margin," which refers to the lower rate of email responses they receive compared to native parents. Next, we investigate whether

Table 2: Treatment Effects on Response Rate

	(1)	(2)	(3)	(4)
Migrant treatment	-0.044***	-0.044***	-0.054***	-0.053***
	(0.007)	(0.007)	(0.010)	(0.010)
Migrant treatment × Higher edu.			0.019	0.018
			(0.014)	(0.014)
Higher education signal			-0.003	-0.004
			(0.009)	(0.009)
Controls	No	Yes	No	Yes
Control group mean (Native sender)	0.707	0.707	0.707	0.707
Scaled treatment effect	-6.3	-6.2	-7.6	-7.5
N	18,663	18,663	18,663	18,663

Notes: Table shows treatment effects on an indicator for whether or not a child care center responds to the email, based on multivariate OLS regressions. *Migrant treatment* is a dummy variable taking a value of one if the email sender's name signals a migration background, and zero if the email sender's name signals a native background. *Higher education signal* is a dummy variable taking a value of one if the email includes a signature that indicates a higher educational background of the sender, and zero if the email does not include a signature. *Controls* include strata fixed effects, additional randomly assigned attributes of the emails (child and sender gender), as well as characteristics of the contacted child care center and the municipality where it is located (see Section 4.1 for details). *Scaled treatment effect* expresses the treatment effect relative to the mean of the respective outcome in the control group of native senders in percent. Robust standard errors in parentheses. Significance levels: * $p < .10$, ** $p < .05$, *** $p < .01$. We additionally report p-values based on randomization inference and correcting for multiple hypothesis testing in Table B2.

discrimination also exists at the "intensive margin" by examining the content of the email responses.

5.2. Results on Email Content

All email content outcomes considered in this section are based on manual ratings of each email response by five independent reviewers who were blind to treatment (see Section 4.2 for details). We estimate treatment effects on six binary dimensions of email content: (i) whether parents are offered a slot (*Slot Offer*), (ii) whether parents are put on a waiting list (*Waiting List*), (iii) whether parents receive a response of above-median length (*Long Response*), (iv) whether the response is rated as "helpful" for parents (*Helpful*), (v) whether the response is rated as "encouraging" (*Encouraging*), and (vi) whether reviewers would recommend a befriended couple with a young child to apply to the responding center (*Recommend*).

We present our findings for these six content dimensions in Table 3. As the response rate is affected by the migrant treatment, there is non-random selection into response

Table 3: Treatment Effects on Response Content

	(1) Slot Offer	(2) Waiting List	(3) Long Resp.	(4) Helpful	(5) Encouraging	(6) Recommend
Panel A (Unconditional)						
Migrant treatment	-0.011***	-0.043***	-0.065***	-0.035***	-0.028***	-0.063***
	(0.003)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Control mean (Native sender)	0.049	0.566	0.467	0.346	0.157	0.419
Scaled treatment effect	-23.1	-7.5	-14.0	-10.2	-17.5	-15.1
N	18,663	18,663	18,663	18,663	18,663	18,663
Panel B (Conditional)						
Migrant treatment	-0.010**	-0.013*	-0.058***	-0.023**	-0.025***	-0.057***
	(0.004)	(0.007)	(0.009)	(0.007)	(0.007)	(0.009)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Scaled treatment effect	-20.5	-2.3	-12.4	-6.5	-15.7	-13.6
N	12,547	12,547	12,547	12,547	12,547	12,547

Notes: Table shows treatment effects on email content measures, based on multivariate OLS regressions. Outcome variables are defined as follows: Column (1): indicator for whether the contacted child care center offers a child care slot before the next turn cycle (August 2022); Column (2): indicator for whether the contacted child care center offers a spot on the waiting list; Column (3): indicator for whether the length of the email response, measured as the number of characters in the email body, is above median; Columns (4) and (5): indicators for whether a child care center responds in a "helpful" or "encouraging" manner; Column (6): indicator for whether the reviewers would recommend the child care center to a befriended couple with a young child. In Panel A, outcome variables receive a value of zero for non-responses (e.g., a non-response is coded as no offer in Column (1)). In Panel B, non-responses are excluded from the estimation sample, so results are conditional on receiving a response. See Section 4.2 and Appendix F for a description of the email rating procedure. *Migrant treatment* is a dummy variable taking a value of one if the email sender's name signals a migration background, and zero if the email sender's name signals a native background. *Controls* include strata fixed effects, additional randomly assigned attributes of the emails (child and sender gender), as well as characteristics of the contacted child care center and the municipality where it is located (see Section 4.1 for details). *Scaled treatment effect* expresses the treatment effect relative to the sample mean (unconditional on response) of the respective outcome in the native control group in percent. Robust standard errors in parentheses. Significance levels: * $p < .10$, ** $p < .05$, *** $p < .01$. We additionally report p-values based on randomization inference and correcting for multiple hypothesis testing in Table B2.

(see Section 4.2 for a discussion). We therefore present two treatment effect estimates for each content outcome: one unconditional on response, including all child care centers that we contacted, where content outcomes of non-responders are coded as zero; and one conditional on response, including only those centers that responded, where content outcomes of non-responders are coded as missing.

Panel A of Table 3 shows that there is a strong negative effect of the migrant treatment on all six dimensions of email content in the full sample (i.e., unconditional on response). All treatment effects are significant at the 1%-level. In particular, parents with a migrant name are 1.1 pp less likely to receive a slot offer, which represents a treatment effect of -23.1% when scaled by the control-group mean. Migrant parents are also 4.3 pp less likely to be put on a waiting list and 6.5 pp less likely to receive a response with above-median

Figure 3: Scaled Treatment Effects on Response Content

Notes: Figure shows scaled treatment effects for email content outcomes, based on the multivariate OLS models shown in Table 3. *Scaled treatment effect* expresses the treatment effect relative to the sample mean (unconditional on response) of the respective outcome in the native control group.

length, which corresponds to treatment effects of -7.5% and -14%, respectively.[19] Email responses to migrant parents are also less positive in tone: They are 3.5 pp (10.2%) less likely to be rated as helpful, 2.8 pp (17.5%) less likely to be rated as encouraging, and 6.3 pp (15.1%) less likely to induce raters to recommend applying to the responding center.

Panel B confirms that all treatment effects also hold in the subsample of centers that responded (i.e., conditional on response). Mechanically, all effects get smaller in size but remain significant at the 5%-level or better, except for the effect on being put on a waiting list ($p = .074$). Thus, migrant parents not only receive fewer responses, but the responses they do receive are also substantially worse in terms of their content.

Next, we examine whether discrimination in content outcomes results from migrant parents not receiving a response (extensive margin) or receiving a worse response (intensive margin). To do so, we express all treatment effects from Table 3 relative to the mean in the native control group (full sample), and compare their magnitude between the

[19]If we estimate the treatment effect on response length measured by the number of characters, we find a treatment effect of -43 characters (or 14.4% shorter emails, $p < .0001$). Given that a German word has on average 5–7 characters, the size of the treatment effect corresponds to emails being about one sentence shorter.

full sample and the subsample of responding centers. The results, displayed in Figure 3, reveal the following key insights: First, all scaled treatment effects (except for *Waiting List*) are large, amounting to at least 10% of the control group mean. Moreover, we observe the largest scaled treatment effects on arguably the most important outcome, as the likelihood of migrant parents receiving a slot offer is more than 20% lower than that of native parents, both unconditional and conditional on response. Second, we find that discrimination in content outcomes is mainly driven by receiving a worse response (i.e., the intensive margin) for all outcomes except for *Waiting List.* For instance, the scaled treatment effect on slot offers is -23.1% unconditional on response (purple bar), and only slightly decreases to -20.5% conditional on response (brown bar).

Overall, our findings demonstrate that child care centers not only discriminate against migrants by not responding to emails, but also through the content of the emails they send in response, which are shorter, less helpful, and less encouraging than those to natives. This is a more subtle and obfuscated form of discrimination than simply not responding to emails (and thus withholding information). In particular, emails to migrants contain substantially fewer slot offers, so child care centers also discriminate in terms of actual enrollment into early child care.

5.3. Exploring the Channels for Discrimination

To be able to address discrimination on the child care market, we need to better understand *why* it occurs. We have already ruled out differential beliefs about the educational background of native and migrant parents as a driver of discrimination (see Section 5.2). In this section, we explore other possible channels for discrimination against migrant parents searching for a child care slot.

To do so, we link our experimental data to a comprehensive set of administrative data from various sources (INKAR, 2021; Bundeswahlleiter, 2021), available at different regional levels (county, constituency, and federal state). We present descriptive evidence on how discrimination varies with regional characteristics to explore possible channels for discrimination.

Contact hypothesis. First, we test the contact hypothesis (Allport, 1954), which posits that direct contact between center managers and migrants should ameliorate discrimination on the child care market. Therefore, we analyze treatment effect heterogeneity with respect to the share of migrant children in child care, measured at the level of the county where the respective child care center is located. In line with the contact hypothesis, we find less discrimination in counties with higher shares of migrant children in child care (see

Table 4: Treatment Effect Heterogeneity on Response Rates

	Response Rate			
	(1)	(2)	(3)	(4)
Migrant treatment	-0.046***	-0.046***	-0.046***	-0.122***
	(0.007)	(0.007)	(0.007)	(0.019)
× Share of migrant children in care (std.)	0.012*			
	(0.007)			
× Right-wing vote share (std.)		-0.020***		
		(0.007)		
× Staff-to-child ratio (std.)			0.019***	
			(0.007)	
× Migrant incentive				0.090***
				(0.020)
Controls	Yes	Yes	Yes	Yes
Observations	17,425	17,412	17,425	18,663

Notes: Table shows treatment effect heterogeneity on the response rate, based on multivariate OLS regressions. Heterogeneity by: Column (1), migrant share in early child care in a county ($n = 400$); Column (2), share of right-wing votes, comprising votes for right-wing parties such as AfD, NPD, and Dritter Weg, measured on a constituency level for the German national election in 2021 ($n = 299$); Column (3), staff-to-child ratio in a county ($n = 400$); Column (4), whether the federal state ($n = 16$) of the child care center has an additional financial incentive for taking up migrant children. The share of migrant children in child care, the right-wing vote share, and the staff-to-child ratio are standardized with mean 0 and standard deviation 1. The smaller number of observations in Columns (1) to (3) is due to missing information about the location of the child care center in the commercial data set (see Appendix E.2). *Migrant treatment* is a dummy variable taking a value of one if the email sender's name signals a migration background, and zero if the email sender's name signals a native background. *Higher education signal* is a dummy variable taking a value of one if the email includes a signature that indicates a higher educational background of the sender, and zero if the email does not include a signature. *Controls* include strata fixed effects, additional randomly assigned attributes of the emails (child and sender gender), as well as characteristics of the contacted child care center and the municipality where it is located (see Section 4.1 for details). Robust standard errors in parentheses. Significance levels: * $p < .10$, ** $p < .05$, *** $p < .01$.

Column (1) of Table 4). The interaction term is statistically significant at the 10%-level ($p = .077$).[20] We interpret these results as tentative evidence in line with the contact hypothesis.

Attitudes towards migrants. Second, discrimination against migrants may be driven by negative attitudes towards migrants in general. Due to social desirability bias, such attitudes are difficult to measure directly. Therefore, we proxy attitudes towards migrants

[20] In a similar vein, we estimate treatment effect heterogeneity with respect to the share of migrants in the municipality (our preregistered control variable, see Section 4.1). Although this measure is clearly a less precise proxy for center managers' direct contact with migrant families, we find qualitatively similar results. The interaction term has a coefficient of $b = .011$, just shy of statistical significance ($p = .120$).

using the vote share for right-wing parties in the 2021 federal election in the constituency of the respective child care center.[21] Column (2) of Table 4 presents result for this heterogeneity analysis. As expected, we observe stronger discrimination against migrants in constituencies with a higher right-wing vote share ($p = 0.003$). Thus, a possible explanation for discrimination on the child care market is that child care center managers have more negative attitudes towards migrants, which would constitute a direct way of discrimination (see Bohren et al., 2022). However, such form of direct discrimination would also be present if child care center managers would not themselves have a preferences against migrants, but rather take parental preferences into account. For example, if child care managers in regions with higher right-wing vote shares expect parents of other children to dislike migrants, they may discourage the enrollment of migrant children to accommodate these parental preferences.

In addition, direct discrimination by child care center managers could potentially have spill-over effects on other stages in the application process. Migrant parents might anticipate adverse treatment by child care center managers and adjust their behavior accordingly by making fewer inquiries and submitting fewer applications, resulting in a systemic type of discrimination (see Bohren et al., 2022). Even if migrant parents do apply, their knowledge of existing discrimination might induce stereotype threat, leading to systematically lower-quality applications (Steele and Aronson, 1995).

Resources and incentives. A third possible explanation for discrimination against migrants is that center managers perceive migrant children as costlier to educate (e.g., due to language barriers or additional organizational tasks related to food or other requirements). To test this hypothesis, we analyze whether discrimination against migrants is more prevalent in child care centers with limited resources, seeking to avoid these additional costs. We use administrative data on the staff-to-child ratio in the county of the respective child care center as a proxy for its resource endowment. Indeed, Column (3) of Table 4 shows that centers in counties with a higher staff-to-child ratio discriminate less ($p = 0.006$).

The notion that educating migrant children may be more costly for child care centers is also evident in public financing arrangements. In Germany, nine out of 16 federal states provide additional resources to child care centers to educate migrant children, mainly in the form of additional funding or staff. If the cost channel is relevant, these incentives

[21] As right-wing parties we consider the "Alternative für Deutschland" (AfD), "Nationaldemokratische Partei Deutschlands" (NPD), "Die Rechte", and "Dritter Weg", which often express xenophobic views.

should reduce discrimination against migrants. Our findings support this hypothesis. In Column (4) of Table 4, we observe substantially stronger discrimination in federal states without financial incentives to educate migrant children ($p < 0.0001$).[22] Note that this analysis has to be interpreted cautiously as these incentives only vary at the federal state level. Nevertheless, combined with the results on the staff-to-child ratio above (which varies at the county-level), our evidence suggests that anticipated higher costs to educate migrant children could be a possible reason for discrimination on the child care market. Thus, policymakers may be able to address discrimination in early child care by providing sufficient resources or financial incentives to educate migrant children.

As a robustness check, we replicate the heterogeneity analysis for slot offers. All heterogeneity results are qualitatively similar, and interaction terms are significant at the 5%-level or better (see Table B1).[23] We also explored treatment effect heterogeneities by provider type (i.e., church vs. public providers), urban classification of the county (i.e., rural vs. city), and economic indicators (i.e., unemployment rate and GDP per capita in the county), but find no meaningful heterogeneities in any of these factors.

5.4. Robustness Checks

All our results are robust to randomization inference and different procedures to correct for multiple hypothesis testing (see Table B2). Because response rate is a binary outcome, we show that our regression results hold when using Probit estimations instead of linear probability models (see Table B3). Furthermore, to compare child care centers that are as homogeneous as possible in (unobservable) characteristics, we conduct an additional robustness check by including zip-code level fixed effects, thus comparing child care centers within a single zip-code area ($n = 3,263$). Table B4 shows that our results are robust to using only within-zip-code variation; in fact, treatment effects even get slightly larger for most outcomes.

Furthermore, we test the robustness of our email content outcomes. To create the content measures, we aggregate the five reviewers' ratings from a four-point Likert scale into binary measures (see Appendix F). We then combine the individual binary ratings into one binary indicator for each content dimension by applying simple majority rule (i.e., we use the rating given by at least three reviewers). However, since this simple majority

[22] However, discrimination against migrants is also significant in states with incentives to educate migrant children (-3.2 pp, $p < 0.0001$).

[23] The only exception is the interaction with incentives to educate migrant children, which is not statistically significant at conventional levels ($p = .138$). However, this variable only varies at the state level.

decision rule is somewhat arbitrary, we check if our results hold when using alternative aggregation rules. Table B5 reports the results for content measures when one, two, three, four, or all five reviewers rated an email as a slot offer, waiting list offer, etc. Results are strikingly consistent across aggregation methods: Treatment effects are consistently negative and qualitatively similar in size, especially for the scaled treatment effects. With a few exceptions, all effects are also statistically significant at the 5%-level or better.

In a similar vein, we check whether our results are sensitive to the transformation of the reviewers' ratings elicited on four-point scales into binary outcomes. Therefore, we also use the full variation in the ratings. To to so, we estimate treatment effects on a standardized version of the four-point scale ratings in Table B6.[24] Reassuringly, all estimated treatment effects on the standardized content outcomes are negative, large, and statistically significant at the 1%-level.

Finally, although the reviewers were blind to the experimental design, treatment assignment, and identity of the child care centers and managers, there is still a possibility that they were subject to some form of (systematic) bias in their ratings. To address this concern, we collaborated with computer scientists to develop a pre-trained natural language processing (NLP) BERT model,[25] which classified the content of email responses for the outcomes *Slot Offer* and *Waiting List* (see Appendix G for details). Our results are robust to using this alternative, computational classification of content outcomes (see Table B7). Remarkably, the scaled treatment effect on slot offers is -20% when employing the computational classification approach, mirroring the results obtained using manual coding.

6. Conclusion

We provide the first causal evidence that migrants are discriminated against when searching and applying for a slot in early child care. Using a randomized email correspondence study with a large, nationwide sample of child care centers in Germany, we find that emails from fictitious migrant parents have a 4.4 pp (6.2%) lower chance of receiving a response compared to emails from native parents. Importantly, discrimination is not only present at the extensive margin of receiving an email response, but also at the intensive

[24]We compute the alternative content outcomes by (i) z-standardizing each reviewer's four-point-scale rating, (ii) creating an equally weighted average of the standardized ratings across all five reviewers, and (iii) z-standardizing the averages again.

[25]BERT stands for Bidirectional Encoder Representations from Transformers. See Devlin et al. (2018) for details.

margin of email content. Responses to migrant parents are less likely to contain slot offers or waiting list offers, are shorter, and rated as less helpful, less encouraging, and generally less appealing in substance and tone. In terms of mechanisms, we find that discrimination is not driven by beliefs about parents' educational background. Intriguingly, however, it is more pronounced in areas with lower share of migrant children in child care, higher right-wing vote shares, and lower staff-to-child ratios. From a policy perspective, these heterogeneity results suggest several avenues how to effectively combat discrimination on the early child care market: providing child care centers with additional resources (e.g., for pedagogical personnel) and financial incentives to enroll migrant children, which are already in place in some federal states.

Our findings indicate that discriminatory practices of child care centers contribute to the large gap in early child care enrollment rates between migrants and natives. This discrimination reinforces existing inequalities of opportunities for disadvantaged children, who could greatly benefit from attending child care in both the short and long term (Cornelissen et al., 2018; Felfe and Lalive, 2018; García et al., 2020). However, discrimination prevents them from realizing these benefits. Importantly, the adverse effects of discrimination on the child care market are likely to extend beyond children themselves. In particular, reemployment opportunities of mothers of migrant children may suffer, as access to child care is an important prerequisite for integration into the labor market (see, e.g., Bauernschuster and Schlotter, 2015; Gambaro et al., 2021; Hermes et al., 2022). The far-reaching consequences of limited child care access for children, mothers, and society at large highlight the need to expand child care capacities and implement structural improvements that promote equitable access to universal early child care.

References

Alan, S., E. Duysak, E. Kubilay, and I. Mumcu (2021). Social Exclusion and Ethnic Segregation in Schools: The Role of Teacher's Ethnic Prejudice. *The Review of Economics and Statistics*, 1–45.

Alesina, A., M. Carlana, E. La Ferrara, and P. Pinotti (2018). Revealing Stereotypes: Evidence from Immigrants in Schools. Working Paper w25333, National Bureau of Economic Research.

Allport, G. W. (1954). *The Nature of Prejudice.* Addison-Wesley.

Alt, C., B. Gedon, S. Hubert, K. Hüsken, and K. Lippert (2019). DJI-Kinderbetreuungsreport 2018 - Inanspruchnahme und Bedarfe bei Kindern bis 14 Jahren aus Elternperspektive - ein Bundesländervergleich. DJI 2019, Deutsches Jugendinstitut, München.

Arcidiacono, P., J. Kinsler, and T. Ransom (2022). Asian American Discrimination in Harvard Admissions. *European Economic Review 144*, 104079.

Baert, S. (2018). Hiring Discrimination: An Overview of (Almost) All Correspondence Experiments Since 2005. In S. M. Gaddis (Ed.), *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*, pp. 63–77. Springer International Publishing.

Bartoš, V., M. Bauer, J. Chytilová, and F. Matějka (2016). Attention Discrimination: Theory and Field Experiments with Monitoring Information Acquisition. *American Economic Review 106*(6), 1437–75.

Bauernschuster, S. and M. Schlotter (2015). Public Child Care and Mothers' Labor Supply—Evidence from Two Quasi-Experiments. *Journal of Public Economics 123*, 1–16.

Bergman, P. and I. McFarlin (2018). Education for All? A Nationwide Audit Study of School Choice. Working Paper w25396, National Bureau of Economic Research.

Bertrand, M. and E. Duflo (2017). Field Experiments on Discrimination. *Handbook of Economic Field Experiments 1*, 309–393.

Bertrand, M. and S. Mullainathan (2004). Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *American Economic Review 94*(4), 991–1013.

Bildungsberichterstattung, A. (2020). *Bildung in Deutschland 2020 - Ein indikatorengestützter Bericht mit einer Analyse zu Bildung in einer digitalisierten Welt.* Bielefeld: wbv Media GmbH.

Bjoerklund, A. and K. G. Salvanes (2011). Education and Family Background: Mechanisms and Policies. In E. A. Hanushek, S. Machin, and L. Woessmann (Eds.), *Handbook of the Economics of Education*, Volume 3, pp. 201–247. North Holland, Amsterdam.

Bohren, J. A., P. Hull, and A. Imas (2022). Systemic Discrimination: Theory and Measurement. Working Paper w29820, National Bureau of Economic Research.

Bundeswahlleiter, D. (2021). Ergebnisse der repräsentativen Wahlstatistik. Statistiken zur Bundestagswahl, Der Bundeswahlleiter, Wiesbaden.

Carlana, M. (2019). Implicit Stereotypes: Evidence from Teachers' Gender Bias. *The Quarterly Journal of Economics 134*(3), 1163–1224.

Cornelissen, T., C. Dustmann, A. Raute, and U. Schoenberg (2018). Who Benefits from Universal Child Care? Estimating Marginal Returns to Early Child Care Attendance. *Journal of Political Economy 126*(6), 2356–2409.

Currie, J. and D. Almond (2011). Human Capital Development Before Age Five. In *Handbook of Labor Economics*, Volume 4, pp. 1315–1486. Elsevier.

de Lafuente, D. M. (2021). Cultural Assimilation and Ethnic Discrimination: An Audit Study with Schools. *Labour Economics 72*, 102058.

Destatis (2020). *Statistiken der Kinder- und Jugendhilfe.* German Federal Statistical Office, Wiesbaden.

Devlin, J., M.-W. Chang, K. Lee, and K. Toutanova (2018). BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding. *arXiv*, preprint:1810.04805.

DJI (2021). Fachkräftebarometer Frühe Bildung 2021. Weiterbildungsinitiative Frühpädagogische Fachkräfte, Autorengruppe Fachkräftebarometer, München.

Drange, N. and T. Havnes (2019). Early Childcare and Cognitive Development: Evidence from an Assignment Lottery. *Journal of Labor Economics 37*(2), 581–620.

Dynarski, S., C. Libassi, K. Michelmore, and S. Owen (2018). Closing the Gap: The Effect of a Targeted, Tuition-Free Promise on College Choices of High-Achieving, Low-Income Students. Working Paper w25349, National Bureau of Economic Research.

Felfe, C. and R. Lalive (2018). Does Early Child Care Affect Children's Development? *Journal of Public Economics 159*, 33–53.

Gambaro, L., G. Neidhöfer, and C. K. Spiess (2021). The Effect of Early Childhood Education and Care Services on the Integration of Refugee Families. *Labour Economics 72*, 102053.

García, J. L., J. J. Heckman, D. E. Leaf, and M. J. Prados (2020). Quantifying the life-cycle benefits of an influential early-childhood program. *Journal of Political Economy 128*(7), 2502–2541.

Giulietti, C., M. Tonin, and M. Vlassopoulos (2019). Racial Discrimination in Local Public Services: A Field Experiment in the United States. *Journal of the European Economic Association 17*(1), 165–204.

Heckman, J., R. Pinto, and P. Savelyev (2013). Understanding the Mechanisms Through which an Influential Early Childhood Program Boosted Adult Outcomes. *American Economic Review 103*(6), 2052–86.

Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Yavitz (2010). The Rate of Return to the High Scope Perry Preschool Program. *Journal of Public Economics 94*(1-2), 114–128.

Hemker, J. and A. Rink (2017). Multiple Dimensions of Bureaucratic Discrimination: Evidence from German Welfare Offices. *American Journal of Political Science 61*(4), 786–803.

Hermes, H., M. Krauß, P. Lergetporer, F. Peter, and S. Wiederhold (2022). Early Child Care and Labor Supply of Lower-SES Mothers: A Randomized Controlled Trial. Working Paper No. 10178, CESifo.

Hermes, H., P. Lergetporer, F. Mierisch, F. Peter, and S. Wiederhold (2023). Males Should Mail? Gender Discrimination in Access to Childcarel. *American Economic Association Papers & Proceedings*, forthcoming.

Hermes, H., P. Lergetporer, F. Peter, and S. Wiederhold (2021). Behavioral Barriers and the Socioeconomic Gap in Child Care Enrollment. Discussion Paper No. 16501, Center for Economic Policy Research.

HRK (2021). Statistische Daten zu Studienangeboten an Hochschulen in Deutschland. Statistiken zur Hochschulpolitik, Hochschulrektorenkonferenz, Berlin.

Hussar, B., J. Zhang, S. Hein, K. Wang, A. Roberts, J. Cui, M. Smith, F. B. Mann, A. Barmer, and R. Dilig (2020). The Condition of Education 2020. *National Center for Education Statistics (NCES 2020-144)*.

INKAR (2021). Indikatoren und Karten zur Raum- und Stadtentwicklung. Statistiken zur räumlichen Zusammensetzung, Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR) im Bundesamt für Bauwesen und Raumordnung (BBR), Bonn.

Jessen, J., S. Schmitz, and S. Waights (2020). Understanding Day Care Enrolment Gaps. *Journal of Public Economics 190*, 104252.

Kline, P., E. K. Rose, and C. R. Walters (2022). Systemic Discrimination Among Large US Employers. *The Quarterly Journal of Economics 137*(4), 1963–2036.

Lavy, V., E. Sand, and M. Shayo (2022). Discrimination Between Religious and Non-Religious Groups: Evidence from Marking High-Stakes Exams. *The Economic Journal 132*, 2308–2324.

Mocan, N. (2007). Can Consumers Detect Lemons? An Empirical Analysis of Information Asymmetry in the Market for Child Care. *Journal of Population Economics 20*, 743–780.

Neumark, D. (2018). Experimental Research on Labor Market Discrimination. *Journal of Economic Literature 56*(3), 799–866.

OECD (2018). Settling In 2018. Technical Report, OECD, Paris.

OECD (2020). Is Childcare Affordable? Policy Brief on Employment, Labour and Social Affairs. Technical Report, OECD, Paris.

Schmitz, S., C. K. Spieß, and M. Huebener (2023). Weiterhin Ungleichheiten bei der KiTa-Nutzung: Größter ungedeckter Bedarf in grundsätzlich benachteiligten Familien. *Bevölkerungsforschung Aktuell 2*, 3–8.

Spiess, C. K. (2008). Early Childhood Education and Care in Germany: The Status Quo and Reform Proposals. *Zeitschrift für Betriebswirtschaftslehre 2008 67*, 1–20.

Spiess, C. K. (2013). Investments in Education: The Early Years Offer Great Potential. *DIW Economic Bulletin 3*(10), 3–10.

Steele, C. M. and J. Aronson (1995). Stereotype Threat and the Intellectual Test Performance of African Americans. *Journal of personality and social psychology 69*(5), 797.

Online Appendix

for

"Discrimination on the Child Care Market:
A Nationwide Field Experiment"

by

Henning Hermes, Philipp Lergetporer, Fabian Mierisch,
Frauke Peter & Simon Wiederhold

Appendix A. Details on Sample and Treatment

Figure A1: Example for Email: Message on Behalf of a Son with Migration Background, No Higher Education Signal

Dear Sir or Madam,

We are looking for a child care slot for our son starting in January 2022.
He is now 1 year and 5 months old.

Do you still have a slot open? How can we apply for a slot?

Thank you!

Sincerely,
Eylül Yildirim

Figure A2: Example for Email: Message on Behalf of a Daughter with Native Background, with Higher Education Signal

Dear Sir or Madam,

We are looking for a child care slot for our daughter starting in January 2022.
She is now 1 year and 5 months old.

Do you still have a slot open? How can we apply for a slot?

Thank you!

Sincerely,
Sebastian Müller

Sebastian Müller, Bachelor of Arts (FH)
Email: Sebastian.Müller0528@gmail.com

Table A1: Cross-Country Comparison of Early Child Care Systems

Country	Migrant-native Enrollment Gap	Reduced Fees	Slots Rationed	Decentralized
France	Yes[a1]	Yes[b1]	Yes[c1]	Yes[d1]
Germany	Yes[a2]	Yes[b2]	Yes[c2]	Yes[d2]
Italy	Yes[a3]	Yes[b3]	Yes[c3]	No[d3]
Spain	Yes[a4]	Yes[b4]	Yes[c4]	No[d4]
UK	Yes[a5]	Yes[b5]	Yes[c5]	Yes[d5]
US	Yes[a6]	Yes[b6]	Yes[c6]	Yes[d6]

Notes: Table shows features of early child care systems of the five largest European countries (in terms of GDP) and the United States, in alphabetical order. Migrant-native Enrollment Gap: "Yes" if children with a migration background are under-represented in early child care. Reduced fees: "Yes" if lower-income families are eligible for fee reductions or exemptions. Slots rationed: "Yes" if average demand for a child care slot exceeds average supply. Decentralized admission decision: "Yes‘" if admission decisions are taken by individual child care centers.

Sources: [a1,a3,a4,a5] OECD (2018); [a2,c2,d2]Jessen et al. (2020); [a6] Cui et al. (2021); [b1,c1,b3,c3,b4,c4,b5,c5] Eurydice (2019); [b2] Felfe and Lalive (2018); [b6] OECD (2020); [c6] Malik et al. (2018); [d1] Expat (2022); [c3] Del Boca et al. (2016); [d4] Harvey (2022); [d5] Renfrewshire Council (2023); [d6] NYC Department of Education (2023).

Table A2: Balancing (Analysis Sample)

	Native (Control)				Migrant (Treatment)			
	Higher edu. = 0		Higher edu. = 1		Higher edu. = 0		Higher edu. = 1	
	Mean	SD	Diff (mean)	*p*-Value	Diff (mean)	*p*-Value	Diff (mean)	*p*-Value
Sender characteristics								
Sender male	0.50	0.50	0.00	0.779	0.00	0.926	-0.01	0.605
Child male	0.50	0.50	-0.00	0.977	0.00	0.876	0.01	0.448
Child care center characteristics								
Center's maximum capacity	68.38	40.09	0.95	0.291	0.23	0.810	-0.51	0.571
Kindergarten (age 3–6 years)	0.93	0.25	0.00	0.465	-0.00	0.856	-0.01	0.120
Daycare (age >6 years)	0.09	0.29	-0.00	0.994	0.00	0.573	0.00	0.472
Provider								
Church	0.25	0.43	0.00	0.872	0.00	0.957	-0.01	0.410
Public	0.18	0.38	-0.00	0.896	-0.00	0.909	-0.00	0.928
Else	0.57	0.49	-0.00	0.969	0.00	0.967	0.01	0.430
Regional Characteristics								
Urban class								
City	0.44	0.50	-0.00	0.812	-0.00	0.939	0.00	0.734
Intermediate	0.37	0.48	0.00	0.879	0.00	0.900	0.00	0.911
Rural	0.18	0.39	0.00	0.908	-0.00	0.953	-0.00	0.563
Share of migrants (in percent)	23.61	11.94	-0.32	0.199	-0.12	0.617	0.18	0.453
Share of migrant children in care (in percent)	29.84	13.03	-0.22	0.403	-0.08	0.765	-0.07	0.796
Staff-to-child ratio	0.13	0.024	0.01	0.451	0.00	0.667	0.00	0.593
Right-wing vote share (in percent)	9.76	5.78	0.09	0.461	0.04	0.709	0.09	0.476
Incentive for migrant children	0.87	0.33	-0.01	0.485	0.02	0.842	-0.05	0.769
State								
Baden Wurttemberg	0.15	0.36	-0.01	0.496	0.00	0.973	0.00	0.573
Bavaria	0.14	0.35	0.00	0.922	-0.01	0.339	-0.00	0.588
Berlin	0.07	0.25	-0.00	0.630	-0.00	0.463	-0.00	0.651
Brandenburg	0.02	0.15	0.00	0.234	-0.00	0.973	-0.00	0.541
Bremen	0.01	0.08	-0.00	0.938	-0.00	0.629	0.00	0.848
Hamburg	0.03	0.17	0.00	0.925	-0.00	0.791	-0.00	0.867
Hesse	0.06	0.24	-0.00	0.643	-0.00	0.617	0.00	0.568
Mecklenburg-Western Pomerania	0.01	0.11	-0.00	0.822	-0.00	0.650	-0.00	0.948
Lower Saxony	0.08	0.27	-0.00	0.914	0.00	0.859	0.00	0.756
North Rhine-Westphalia	0.23	0.42	0.00	0.625	0.01	0.234	-0.00	0.831
Rhineland-Palatinate	0.04	0.19	-0.00	0.787	0.00	0.534	0.00	0.641
Saarland	0.01	0.12	0.00	0.621	0.00	0.710	-0.00	0.940
Saxony	0.06	0.23	0.00	0.477	0.00	0.713	0.00	0.910
Saxony-Anhalt	0.01	0.11	-0.00	0.546	-0.00	0.946	0.00	0.750
Schleswig-Holstein	0.04	0.19	-0.00	0.656	-0.00	0.692	-0.00	0.777
Thuringia	0.02	0.15	0.00	0.678	0.00	0.815	-0.00	0.974
Total (N=18,663)	4,682		4,631		4,661		4,689	
Bounces (N= 3,795)	941		974		956		924	
Sent (N=22,458)	5,623		5,605		5,617		5,613	

Notes: Table shows means and standard deviations of variables by treatment group. The analysis sample excludes "bounced" emails. *Migrant treatment* is a dummy variable taking a value of one if the email sender's name signals a migration background, and zero if the email sender's name signals a native background. *Higher edu.* is a dummy variable taking a value of one if the email includes a signature that indicates a higher educational background of the sender, and zero if the email does not include a signature. For variable definitions, see Appendix E. *Diff* is the difference in the mean of the respective variable between the baseline group (native treatment and no higher education signal) and each of the other three experimental groups. We report p-values for two-sided t-tests of the null hypothesis that differences are equal to zero. For variables we use in the heterogeneity analysis we impute missing values by the mean of a higher regional level (municipality or county). Significance levels: * $p < .10$, ** $p < .05$, *** $p < .01$.

Table A3: Balancing (Sent Emails)

	Native (Control)				Migrant (Treatment)			
	Higher edu. = 0		Higher edu. = 1		Higher edu. = 0		Higher edu. = 1	
	Mean	SD	Diff (mean)	p-Value	Diff (mean)	p-Value	Diff (mean)	p-Value
Sender characteristics								
Sender male	0.50	0.50	0.00	0.720	0.00	0.865	-0.00	0.791
Child male	0.50	0.50	-0.00	0.881	-0.00	0.807	0.00	0.665
Child care center characteristics								
Center's maximum capacity	67.98	39.70	0.93	0.243	0.28	0.734	-0.53	0.507
Kindergarten (age 3–6 years)	0.93	0.25	0.00	0.521	-0.00	0.868	-0.01	0.119
Daycare (age >6 years)	0.10	0.29	-0.00	0.534	0.00	0.985	0.00	0.435
Provider								
Church	0.24	0.43	0.00	0.977	-0.00	0.920	-0.00	0.896
Public	0.17	0.38	-0.00	0.941	-0.00	0.980	0.00	0.927
Else	0.59	0.49	0.00	0.975	0.00	0.915	0.00	0.965
Regional characteristics								
Urban class								
City	0.43	0.49	0.00	0.839	0.00	0.901	-0.00	0.918
Intermediate	0.38	0.49	-0.00	0.936	-0.00	0.912	0.00	0.939
Rural	0.19	0.39	-0.00	0.875	-0.00	0.984	0.00	0.972
Share of migrants (in percent)	23.49	11.89	-0.09	0.676	-0.13	0.548	-0.04	0.843
Share of migrant children in care (in percent)	29.77	13.02	0.18	0.944	-0.12	0.623	-0.03	0.917
Staff-to-child ratio	0.13	0.024	0.00	0.878	0.00	0.453	0.00	0.911
Right-wing vote share (in percent)	9.81	5.69	-0.02	0.882	0.00	0.995	0.09	0.418
Incentive for migrant children	0.88	0.33	-0.02	0.800	-0.01	0.971	-0.02	0.976
State								
Baden Wurttemberg	0.16	0.37	-0.00	0.937	-0.00	0.938	0.00	0.904
Bavaria	0.15	0.36	0.00	0.984	-0.00	0.998	-0.00	0.995
Berlin	0.06	0.24	0.00	0.903	0.00	0.957	-0.00	0.953
Brandenburg	0.03	0.16	0.00	0.930	0.00	0.945	-0.00	0.943
Bremen	0.01	0.08	-0.00	0.921	-0.00	0.914	-0.00	0.995
Hamburg	0.03	0.16	0.00	0.977	0.00	0.947	-0.00	0.990
Hesse	0.06	0.24	-0.00	0.941	-0.00	0.918	0.00	0.953
Mecklenburg-Western Pomerania	0.01	0.11	-0.00	0.676	-0.00	0.868	0.00	0.936
Lower Saxony	0.08	0.27	0.00	0.960	0.00	0.959	0.00	0.963
North Rhine-Westphalia	0.23	0.42	0.00	0.838	0.00	0.834	-0.00	0.878
Rhineland-Palatinate	0.04	0.20	-0.00	0.915	-0.00	0.896	0.00	0.862
Saarland	0.01	0.12	0.00	0.984	0.00	0.867	-0.00	0.928
Saxony	0.05	0.22	-0.00	0.965	-0.00	0.944	-0.00	0.952
Saxony-Anhalt	0.01	0.11	-0.00	0.740	-0.00	0.798	0.00	0.866
Schleswig-Holstein	0.03	0.18	0.00	0.892	-0.00	0.967	-0.00	0.907
Thuringia	0.02	0.15	-0.00	0.924	0.00	0.993	0.00	0.961
Bounces (N= 3,795)	0.17	0.37	0.01	0.365	0.00	0.687	-0.01	0.196
Sent (N= 22,458)	5,623		5,605		5,617		5,613	

Notes: Table shows means and standard deviations of variables by treatment group. The sent sample includes "bounced" emails. *Migrant treatment* is a dummy variable taking a value of one if the email sender's name signals a migration background, and zero if the email sender's name signals a native background. *Higher edu.* is a dummy variable taking a value of one if the email includes a signature that indicates a higher educational background of the sender, and zero if the email does not include a signature. For variable definitions, see Appendix E. *Diff* is the difference in the mean of the respective variable between the baseline group (native treatment and no higher education signal) and each of the other three experimental groups. We report p-values for two-sided t-tests of the null hypothesis that differences are equal to zero. For variables we use in the heterogeneity analysis we impute missing values by the mean of a higher regional level (municipality or county). Significance levels: * $p < .10$, ** $p < .05$, *** $p < .01$.

Appendix B. Robustness Checks

Table B1: Treatment Effect Heterogeneity on Slot Offers

	Slot Offer			
	(1)	(2)	(3)	(4)
Migrant treatment	-0.012***	-0.012***	-0.012***	-0.027**
	(0.003)	(0.003)	(0.003)	(0.012)
× Share of migrant children in care (std.)	0.009**			
	(0.004)			
× Right-wing vote share (std.)		-0.015***		
		(0.005)		
× Staff-to-child ratio (std.)			0.009**	
			(0.004)	
× Migrant incentive				0.018
				(0.012)
Controls	Yes	Yes	Yes	Yes
Observations	17,425	17,412	17,425	18,663

Notes: Table shows treatment effect heterogeneity on slot offers, based on multivariate OLS regressions. Heterogeneity by: Column (1), migrant share in early child care in a county; Column (2), share of right-wing votes, comprising votes for right-wing parties such as AfD, NPD, and Dritter Weg, measured on a constituency level for the German national election in 2021; Column (3), staff-to-child ratio in a county; Column (4), whether the federal state of the child care center has an additional financial incentive for taking up migrant children. The share of migrant children in child care, the right-wing vote share, and the staff-to-child ratio are standardized with mean 0 and standard deviation 1. Missing observations are due to incomplete information on the location of the child care center in the commercial data set. *Migrant treatment* is a dummy variable taking a value of one if the email sender's name signals a migration background, and zero if the email sender's name signals a native background. *Higher education signal* is a dummy variable taking a value of one if the email includes a signature that indicates a higher educational background of the sender, and zero if the email does not include a signature. *Controls* include strata fixed effects, additional randomly assigned attributes of the emails (child and sender gender), as well as characteristics of the contacted child care center and the municipality where it is located (see Section 4.1 for details). Robust standard errors in parentheses. Significance levels: * $p <$ .10, ** $p <$ .05, *** $p <$ .01.

Table B2: Randomization Inference and Corrections for Multiple Hypothesis Testing

	Coefficient (1)	p-value Main Table (2)	Rand. Inference (3)	List-Shakih-Xu (4)	Westphal-Young (5)	Romano-Wolf (6)
Panel A: Migrant and Higher Education Signal (Table 2)						
Migrant treatment	-0.044***	**0.000**	**0.000**	**0.001**	**0.000**	**0.001**
Migrant treatment × Higher edu.	0.018	0.203	0.181	0.216	0.203	0.221
Panel B: Content Outcomes - Unconditional (Table 3, Panel A)						
Slot Offer	-0.011***	**0.000**	**0.000**	**0.001**	**0.000**	**0.001**
Waiting list	-0.043***	**0.000**	**0.000**	**0.001**	**0.000**	**0.001**
Long response	-0.065***	**0.000**	**0.000**	**0.001**	**0.000**	**0.001**
Helpful	-0.035***	**0.000**	**0.000**	**0.001**	**0.000**	**0.001**
Encouraging	-0.028***	**0.000**	**0.000**	**0.001**	**0.000**	**0.001**
Recommend	-0.063***	**0.000**	**0.000**	**0.001**	**0.000**	**0.001**
Panel C: Content Outcomes - Conditional (Table 3, Panel B)						
Offer	-0.010**	**0.018**	**0.012**	**0.017**	**0.018**	**0.018**
Waiting list	-0.013*	**0.073**	**0.061**	**0.065**	**0.074**	**0.073**
Long response	-0.058***	**0.000**	**0.000**	**0.001**	**0.000**	**0.001**
Helpful	-0.023**	**0.013**	**0.013**	**0.012**	**0.013**	**0.014**
Encouraging	-0.025***	**0.001**	**0.000**	**0.002**	**0.001**	**0.002**
Recommend	-0.057***	**0.000**	**0.000**	**0.001**	**0.000**	**0.001**

Notes: Table shows p-values for our main results when using randomization inference and adjusting for multiple hypothesis testing. All p-values < .10 are printed in bold. For comparison, Column (1) displays coefficients and significance stars representing p-values from robust standard errors (* $p < .10$, ** $p < .05$, *** $p < .01$) as reported in the main tables. Column (2) shows the p-values as reported in the main tables. Randomization inference (RI) p-values in Column (3) are obtained from RI with 1,000 permutations, assigning the treatment status randomly within strata (using the Stata command 'ritest' by Heß (2017)). In Columns (4)—(6), we implement three different methods to correct for multiple hypothesis testing (controlling the family-wise error rates) using bootstrap resampling techniques. Column (4) uses the method by List et al. (2019), Column (5) the stepdown-approach by Westfall and Young (1993), and Column (6) the approach by Romano and Wolf (2005, 2016). The procedures by Westfall-Young (using the Stata command 'wyoung' by Julian Reif) and Romano-Wolf (using the Stata command 'rwolf' by Clarke et al. (2020)) account for the stratified randomization, that is, bootstrap samples are selected within each stratum. In Panel A, we correct for the fact that we have two treatments, migrant background and higher education signal. In Panel B (Panel C), we correct for the multiple email content dimensions unconditional (conditional) on response. Note that some corrected p-values are smaller than the original p-values because they are based on bootstrap methods. All control variables from the respective baseline specification are included.

Table B3: Treatment Effects on Response Rate and Content Outcomes (Probit Regression)

	(1) Response Rate	(2) Slot Offer	(3) Waiting List	(4) Long Response	(5) Helpful	(6) Encouraging	(7) Recommend
Migrant treatment	-0.127***	-0.137***	-0.111***	-0.159***	-0.102***	-0.125***	-0.171***
	(0.020)	(0.036)	(0.019)	(0.024)	(0.020)	(0.024)	(0.020)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Marginal effect	-0.044	-0.011	-0.043	-0.058	-0.035	-0.026	-0.063
N	18,617	17,678	18,652	12,529	18,633	18,623	18,634

Notes: Table shows treatment effects based on multivariate Probit regressions. Outcome variables are defined as follows: Column (1): indicator for whether or not a child care center responds to the email; Column (2): indicator for whether the contacted child care center offers a child care slot before the next turn cycle (August 2022); Column (3): indicator for whether the contacted child care center offers a place on the waiting list; Column (4): indicator for whether the length of the email response, measured as the number of characters in the email body, is above median; Columns (5) and (6): indicators for whether a child care center responds in a "helpful" or "encouraging" manner; Column (7): indicator for whether the reviewers would recommend the child care center to a befriended couple with a young child. Outcome variables receive a value of zero for non-responses (e.g., a non-response is coded as no offer in Column (2)). See Section 4.2 and Appendix F for a description of the email rating procedure. *Migrant treatment* is a dummy variable taking a value of one if the email sender's name signals a migration background, and zero if the email sender's name signals a native background. *Controls* include strata fixed effects, additional randomly assigned attributes of the emails (child and sender gender), as well as characteristics of the contacted child care center and the municipality where it is located (see Section 4.1 for details). Marginal effects are reported in the bottom of the table. Robust standard errors in parentheses. Significance levels: * $p <$ .10, ** $p <$ .05, *** $p <$ .01.

Table B4: Treatment Effects with Zip-Code Fixed Effects

	(1) Response Rate	(2) Slot Offer	(3) Waiting List	(4) Long Response	(5) Helpful	(6) Encouraging	(7) Recommend
Migrant treatment	-0.047***	-0.012***	-0.052***	-0.071***	-0.039***	-0.026***	-0.065***
	(0.008)	(0.003)	(0.009)	(0.009)	(0.008)	(0.006)	(0.008)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control mean (Native sender)	0.707	0.049	0.566	0.467	0.346	0.157	0.419
Scaled treatment effect	-6.7	-23.8	-9.2	-15.1	-11.3	-16.3	-15.6
N	16,917	16,917	16,917	16,917	16,917	16,917	16,917

Notes: Table shows treatment effects on response rate and email content measures, based on multivariate OLS regressions. Outcome variables are defined as follows: Column (1): indicator for whether or not a child care center responds to the email; Column (2): indicator for whether the contacted child care center offers a child care slot before the next turn cycle (August 2022); Column (3): indicator for whether the contacted child care center offers a place on the waiting list; Column (4): indicator for whether the length of the email response, measured as the number of characters in the email body, is above median; Columns (5) and (6): indicators for whether a child care center responds in a "helpful" or "encouraging" manner; Column (7): indicator for whether the reviewers would recommend the child care center to a befriended couple with a young child. All outcome variables receive a value of zero for non-responses (e.g., a non-response is coded as no offer in Column (1)). Scaled effects are calculated using the respective outcome mean in the native control group. *Migrant treatment* is a dummy variable taking a value of one if the email sender's name signals a migration background, and zero if the email sender's name signals a native background. *Controls* include strata fixed effects, additional randomly assigned attributes of the emails (child and sender gender), as well as characteristics of the contacted child care center and the municipality where it is located (see Section 4.1 for details). Controls also include zip-code fixed effects. Regressions leave out 1,746 observations without treatment variation within a zip-code area. Estimating our preregistered model on this smaller sample yields very similar results as in the full sample in Tables 2 and 3. Robust standard errors in parentheses. Significance levels: * $p < .10$, ** $p < .05$, *** $p < .01$.

Table B5: Treatment Effects for Different Definitions of Content Outcomes

	1 Reviewer		2 Reviewers		3 Reviewers		4 Reviewers		5 Reviewers	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Offer										
Migrant treatment	-0.024***	-0.023***	-0.014***	-0.013***	-0.011***	-0.010**	-0.009***	-0.007*	-0.006***	-0.005
	(0.004)	(0.006)	(0.003)	(0.005)	(0.003)	(0.004)	(0.003)	(0.004)	(0.002)	(0.003)
Control group mean (Native sender)	0.103	0.103	0.060	0.060	0.049	0.049	0.041	0.041	0.028	0.028
Scaled treatment effect	-22.9	-22.2	-23.3	-21.1	-23.1	-20.5	-21.2	-17.1	-22.3	-17.7
Waiting List										
Migrant treatment	-0.054***	-0.022***	-0.050***	-0.021***	-0.043***	-0.013*	-0.039***	-0.010	-0.042***	-0.018**
	(0.007)	(0.005)	(0.007)	(0.006)	(0.007)	(0.007)	(0.007)	(0.008)	(0.007)	(0.008)
Control group mean (Native sender)	0.650	0.650	0.608	0.608	0.566	0.566	0.537	0.537	0.494	0.494
Scaled treatment effect	-8.3	-3.4	-8.3	-3.4	-7.5	-2.3	-7.3	-1.9	-8.5	-3.7
Helpful										
Migrant treatment	-0.050***	-0.032***	-0.048***	-0.035***	-0.035***	-0.023**	-0.024***	-0.012	-0.012**	-0.002
	(0.007)	(0.009)	(0.007)	(0.009)	(0.007)	(0.009)	(0.007)	(0.009)	(0.006)	(0.008)
Control mean (Native sender)	0.483	0.483	0.410	0.410	0.346	0.346	0.276	0.276	0.183	0.183
Scaled treatment effect	-10.4	-6.7	-11.6	-8.4	-10.2	-6.5	-8.8	-4.5	-6.4	-1.2
Encouraging										
Migrant treatment	-0.055***	-0.041***	-0.042***	-0.032***	-0.028***	-0.025***	-0.017***	-0.018***	-0.009***	-0.008**
	(0.007)	(0.009)	(0.007)	(0.009)	(0.005)	(0.007)	(0.003)	(0.005)	(0.003)	(0.004)
Control mean (Native sender)	0.461	0.461	0.312	0.312	0.157	0.157	0.061	0.061	0.038	0.038
Scaled treatment effect	-11.9	-8.8	-13.3	-10.4	-17.5	-15.7	-28.6	-30.0	-24.3	-21.5
Recommend										
Migrant treatment	-0.061***	-0.033***	-0.065***	-0.051***	-0.063***	-0.057***	-0.054***	-0.055***	-0.030***	-0.033***
	(0.007)	(0.006)	(0.007)	(0.008)	(0.007)	(0.009)	(0.006)	(0.009)	(0.005)	(0.007)
Control mean (Native sender)	0.636	0.636	0.521	0.521	0.419	0.419	0.276	0.276	0.124	0.124
Scaled treatment effect	-9.5	-5.2	-12.5	-9.8	-15.1	-13.6	-19.5	-20.1	-23.9	-26.4
Conditional on response	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	18,663	12,547	18,663	12,547	18,663	12,547	18,663	12,547	18,663	12,547

Notes: Table shows treatment effects on email content measures, based on multivariate OLS regressions. Columns (1) and (2) show migrant treatment effects on the respective content outcome when only one of the five reviewers rated the email as an offer, waiting list offer, helpful, encouraging, or recommendable. Analogously, Columns (3)-(10) report results when two, three, four, or all five reviewers considered the email as an offer, waiting list offer, helpful, encouraging, or recommendable. Uneven columns show effects on unconditional on response and even columns show effects on conditional on response. *Scaled treatment effect* expresses the treatment effect relative to the mean of the respective outcome in the control group of native senders in percent. *Migrant treatment* is a dummy variable taking a value of one if the email sender's name signals a migration background, and zero if the email sender's name signals a native background. *Controls* include strata fixed effects, additional randomly assigned attributes of the emails (child and sender gender), as well as characteristics of the contacted child care center and the municipality where it is located (see Section 4.1 for details). Robust standard errors in parentheses. Significance levels: * $p <$ .10, ** $p <$ .05, *** $p <$ .01.

Table B6: Treatment Effects on Standardized Response Content

	Slot Offer		Waiting List		Helpful		Encouraging		Recommend	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Migrant treatment	-0.056***	-0.041**	-0.085***	-0.032*	-0.075***	-0.045**	-0.079***	-0.060***	-0.130***	-0.115***
	(0.015)	(0.017)	(0.015)	(0.018)	(0.015)	(0.018)	(0.015)	(0.018)	(0.015)	(0.018)
Conditional on response	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	18,663	12,547	18,663	12,547	18,663	12,547	18,663	12,547	18,663	12,547

Notes: Table shows treatment effects on email content measures, based on multivariate OLS regressions. Email content outcomes are measured as the mean of the standardized ratings of all five reviewers, standardized with mean 0 and standard deviation 1. Uneven columns show effects on unconditional on response and even columns show effects on conditional on response. *Migrant treatment* is a dummy variable taking a value of one if the email sender's name signals a migration background, and zero if the email sender's name signals a native background. *Controls* include strata fixed effects, additional randomly assigned attributes of the emails (child and sender gender), as well as characteristics of the contacted child care center and the municipality where it is located (see Section 4.1 for details). Robust standard errors in parentheses. Significance levels: * $p < .10$, ** $p < .05$, *** $p < .01$.

Table B7: Treatment Effects for Outcomes Classified by NLP

	Slot Offer (BERT)		**Waiting List (BERT)**	
	(1)	(2)	(3)	(4)
Migrant treatment	-0.014***	-0.014***	-0.045***	-0.016**
	(0.004)	(0.005)	(0.008)	(0.007)
Conditional on response	No	Yes	No	Yes
Controls	Yes	Yes	Yes	Yes
Control Mean (Native Sender)	0.066	0.066	0.565	0.565
Scaled Treatment Effect	-21.9	-21.4	-8.0	-2.9
N	17,563	11,447	17,563	11,447

Notes: Table shows treatment effects (estimated by multivariate OLS regressions) on selected email content outcomes classified by a BERT NLP algorithm. Column (1): indicator for whether the contacted child care center offers a child care slot before the next turn cycle (August 2022); Column (2): indicator for whether the contacted child care center offers a place on the waiting list. The sample excludes all data which is used for training and testing the NLP algorithm (1,100 observations). Uneven columns show effects unconditional on response and even columns show effects conditional on response. *Migrant treatment* is a dummy variable taking a value of one if the email sender's name signals a migration background, and zero if the email sender's name signals a native background. *Controls* include strata fixed effects, additional randomly assigned attributes of the emails (child and sender gender), as well as characteristics of the contacted child care center and the municipality where it is located (see Section 4.1 for details). Robust standard errors in parentheses. Significance levels: * $p < .10$, ** $p < .05$, *** $p < .01$.

Appendix C. Additional Surveys

Appendix C.1. Survey with Child Care Centers

In summer 2019, we conducted a series of interviews with six child care centers to further inform our study design. Based on these interviews, we then conducted a survey with more than 400 child care centers in September 2019. To recruit participants for the survey, we sent invitations to a subset of 6,000 randomly chosen early child care centers from the commercial data set and specifically asked the child care center manager to take part in the survey (these 6,000 centers are excluded from our main study). Importantly, in the present study, we did not contact these centers again. The invitation email specifically stated that only center managers have the necessary information to participate in the survey and should participate in the survey.

We distributed the invitation emails in two waves, spaced two weeks apart from each other, to account for differing German vacation schedules. One week after our invitation and four days after our initial reminder, we then sent two waves of additional reminders to take part in our survey. We ask center managers about the institutional background of the center (provider, size, and age), the demographics of the enrolled children (gender, age, and migration background), the criteria for preferential treatment during the enrollment process, the decision-maker for enrollment decisions, and the center's communication with parents (content of the parental requests and communication device).

In the survey, 85% of managers state that they regularly receive email requests from parents, most frequently asking about open slots and information on how to apply (the two questions we also ask in our email). Furthermore, almost 40% of managers indicate that prior email contact is important for parents to ultimately receive a slot. Thus, emails are an essential communication tool between parents and child care centers. In addition, 86% of child care center managers report that the center manager is responsible for enrollment decisions (using the survey question "Who decides on the admission of a child in your facility?"). In contrast, less than 10% of respondents indicate that the rules of the provider or other institutions (such as other employees of the child care center or the city administration) play a role for these decisions.

Appendix C.2. Online Survey

In April and October 2019, we conducted two surveys with $n = 200$ online workers using Clickworker.de, a German counterpart to Amazon MTurk, where individuals can take short surveys or complete other tasks for micro-payments. We use the survey to validate the origin of the names in our study and to check the realism of our emails. Our inclusion criteria for the survey were that participants had to live in Germany and speak German as their native language.

We use the first survey to pre-select names for our experiment. To this end, survey participants had to state if they associate a migration background with the most common names for native Germans living in Germany and for Turkish migrants living in Germany. Participants repeat this exercise for a total of 16 surnames (eight German, eight Turkish) and 56 first names (14 for each gender and country pair). We retrieve names from the online portal of the Society for German Language (GfdS) to identify the most common names from the cohort born in 1986. Finally, we choose the four Turkish and German surnames as well as the four first names for each gender country pair that are most clearly associated with a native or migration background to be validated for their exact origin in the second survey.

In the second survey, we asked participants to write the country of origin next to each of the names identified in the first survey. In our field experiment, we then include the names that were most clearly associated with a German or Turkish origin (at least 90% of respondents associated each of the chosen surnames with either a German or Turkish origin). We then showed participants the email for the field experiment, including the higher education signal, and asked participants if they think the email is realistic and if they perceived the information conveyed in the signature. Of the $n = 182$ participants that responded to the question, $n = 145$ (80%) see our email as realistic or very realistic and only $n = 8$ (4%) see our email as not or not at all realistic. Importantly, we cannot detect any differences in perceived realism based on the name of the email sender (signaling native or migration background). Moreover, more than 80% of the respondents correctly recalled that the sender has a university degree, based on the higher education signal included in the email signature.

Appendix D. Turkish Migrants in Germany

Turkish immigrants are by far the largest and most regionally dispersed ethnic group in Germany. In 2019, there were approximately 1.5 million people of Turkish origin in Germany. This accounts for approximately 1.3 percent of the German population and 13.0 percent of all migrants in Germany. The second largest group are Polish immigrants (0.9% of the German population; about 7.4% of all migrants). Turkish migrants first came to Germany in the 1960s to expand Germany's labor force. The influx of people of Turkish origin was implemented based on an agreement with Turkey to recruit guest workers. In the 1970s, there was a second wave of migration from Turkey to Germany due to family reunification and political instability in Turkey. As a result, the majority (52,6%) of people with a Turkish migration background are second or third generation migrants with no personal migration experience (Bundesamt für Migration, 2019).

Despite the long history of Turkish migrants in Germany, people with Turkish migration backgrounds are less educated and earn less than native Germans. For instance, in the overall German population, about 7% of Turkish persons hold a university degree, while 17% of persons without migration background hold a university degree. Furthermore, the average monthly income per capita of persons in Germany without a migration background is 1,776 EUR, while persons with a Turkish migration background have an average monthly income per capita of 1,237 EUR (Destatis, 2021). Past evidence also shows that persons with Turkish migration background are discriminated against on the German labor market (Kaas and Manger, 2012).

Most important for our setting, children from Turkish migrants are substantially less often enrolled in early child care than children from German parents (11.7% vs. 33.2%). However, Turkish parents state that they want to enroll their child into early child care almost as often as German parents (40.2% vs. 43.9%, see Jessen et al., 2020).

Appendix E. Data Section

Appendix E.1. Stratification Variables

We use three variables in the stratified randomization: The child care center's provider type, urban type of the county of the center, and the federal state in which the center is located. Information on the provider type is part of the commercial data set we bought. We classify the provider type in three categories. We categorize ecclesiastic providers such as the Catholic and Protestant church as well as providers that are close to the church (e.g., "Diakonie") as "church". All types of public providers are classified as "public", while all other providers (e.g, for-profit providers, clubs, parental initiatives, and missings) are assigned to the "other" category. Information on the provider type is missing for 29% of observations. Data on a county's urban type (referring to the year 2021) are taken from Eurostat (see Eurostat). The measure uses inhabitants per square kilometer to classify three degrees of urbanity of a county (i.e., NUTS3 region), namely "city", "intermediate", and "rural". In a first step, the approach defines "urban clusters" as contiguous grid cells of one square kilometer with a density of at least 300 inhabitants per square kilometer and a joint minimum population of 5,000. All other counties are classified as " rural areas". In a second step, a county is classified as "rural" if the share of the population living in rural areas is higher than 50%, as "intermediate" if the share of the population living in rural areas is between 20% and 50%, and "city" if the share of the population living in rural areas is below 20%. Finally, the classification takes the presence of cities into account. A rural county containing a city of more than 200,000 inhabitants making up at least 25% of the county's population is classified as "intermediate", and an "intermediate" county containing a city of more than 500,000 inhabitants making up at least 25% of the county's population is classified as "city"(see also Dijkstra et al., 2019). The federal state is taken from the commercial data set.

In a few cases, the number of observations within a stratum was small (< 8). We, therefore, decided to merge a total of 27 observations (out of 22,458) with the neighboring strata in terms of urban classification (i.e., we either merged "rural" with "intermediate" or "intermediate" with "city"). In total, we stratify on 131 strata in our randomization.

Appendix E.2. Control Variables

We use the following control variables in our regressions:

- Sender Controls: Parent male equals one for accounts with male first names (Andreas, Sebastian, Hüseyin, and Ömer) and is zero for female parent names (Christina, Stefanie, Eylül, and Fatma); child male equals one when the child mentioned in the email is male ("son"), and is zero for female children ("daughter"); an indicator variable that is equal to one if the email was sent from the account with the spelling mistake, and its interaction with the higher education signal.

- Regional Controls: Share of migrants in a municipality retrieved from Microm data (see Microm webpage). There are 3,707 municipalities in our sample. Data are missing for 1,299 observations; we impute missings by the mean of a higher regional level (if possible NUTS3, otherwise NUTS2), and add an imputation dummy to the regressions.

- Child Care Center Controls: Provider type (see Appendix E.1); Center's maximum capacity, referring to the maximum number of children a child care center is allowed to enroll (data are missing for 1,197 observations; we impute missings by the mean of the next-higher regional level, and add an imputation dummy); a kindergarten dummy, which equals one if the child care center also caters to children aged three to six years, zero otherwise; a daycare dummy, which equals one if the child care center also offers a daycare for schoolchildren, zero otherwise. All child care center controls are taken from the commercial data set.

- Strata Fixed Effects: A binary variable which indicates if the center belongs to one of the 131 strata.

Appendix E.3. Further Variables

We use the following variables for our investigation into the possible mechanisms through which discrimination may occur in Section 5.3 (note that we also include these variables in our balancing tests in Tables A2 and A3):

- Share of migrant children in care: The share of migrant children of all children in child care, measured at the level of the county of the child care center in 2019. There are 400 counties in our sample. The data are retrieved from INKAR (2021). Data are missing for 1,238 observations; hence, we estimate specifications using this variable in a sample of 17,425 observations (see Column (1) of Table 4).

- Right-wing vote shares: The share of party votes ("Zweitstimme") for right-wing parties (AfD, NPD, "Dritter Weg", and "Die Rechte") in the 2021 federal elections, measured at the level of the constituency of the child care center. The constituency refers to a specific geographic area or electoral district that is represented by a member of parliament in the federal government. There are 299 constituencies in our sample. The data are retrieved from the German Statistical Office for Voting Data (Bundeswahlleiter, 2021). Data are missing for 1,251 observations; hence, we estimate specifications using this variable in a sample of 17,412 observations (see Column (2) of Table 4).

- Staff-to-child ratio: The ratio of pedagogical staff to the number of slots in a child care center, measured at the level of the county of the child care center in 2019. The data are retrieved from INKAR (2021). Data are missing for 1,238 observations; hence, we estimate specifications using this variable in a sample of 17,425 observations (see Column (3) of Table 4).

- Migrant incentive: A dummy variable equal to one if the relevant laws of the child care center's federal state do specify any subsidies or other advantages for enrolling children with migration background in child care centers, zero otherwise. There are 16 federal states in our sample. We have no missings for this variable; hence, we estimate specifications using this variable in a sample of all 18,663 observations (see Column (4) of Table 4).

Appendix F. Rating Data

Five reviewers evaluate each response from child care centers on the following five outcome dimensions:

1. Whether the email contains an offer for a slot before August 2022 (*Slot offer*).
2. Whether the email contains an offer to get a place on a waiting list (*Waiting list*).
3. Whether the email is perceived as helpful (*Helpful*).
4. Whether the email is perceived as encouraging (*Encouraging*).
5. Whether one would recommend contacting the child care center for a befriended couple with a child aged 1.5 years (*Recommendation*).

For the rating, we recruited five student assistants ("reviewers") from three different universities in Germany during February 2022. Three of the recruited reviewers were female, two were male. Moreover, two were currently in their Master's studies, while three were studying towards a Bachelor's degree. We trained the reviewers in a four-hour workshop in which we went through the rating criteria and jointly reviewed a set of practice emails. Additionally, we provided a handbook on how to code the different outcome dimensions and also supplied them with a rating tool to help them minimize technical errors while reviewing.

In the rating tool, reviewers answered a short survey for each of the 12,547 responses. Each survey question represented one of the outcome dimensions shown above. The respective email response text was visible at the top of the screen at all points in the survey. The questions had to be answered on a four-point Likert scale (except for the recommendation outcome, which was measured on a ten-point Likert scale) from "Clearly not ..." to "Clearly ...". We provided reviewers with completely anonymous emails, i.e., we deleted names of the center managers and the parents beforehand. Thus, reviewers were unaware of the treatment variation. Furthermore, we did not inform reviewers about the purpose of the study.

To use the reviewer results for our analysis, we create binary outcome measures in two steps. First, for each reviewer, we created binary measures by coding emails rated as "Clearly not ..." or "Somewhat not ..." as not a slot offer/waiting list offer/helpful/encouraging. Similarly, we coded emails rated as "Somewhat ..." or "Clearly ..." as a slot offer/waiting list offer/helpful/encouraging. For the recommendation dimension, we create the binary measure by combining answers between scale points one to five to "no recommendation", and combined answers from scale points six to 10 to "recommendation".

In the second step, we combine the individual ratings into one variable. We code a dimension as a slot offer/waiting list offer/helpful/encouraging/recommendation, if the email was rated by three or more reviewers as such, zero otherwise. Results are robust to different specifications of the binary measure (see Table B5) and reviewer-specific use of the scales (see Table B6).

We compute measures of inter-rater reliability for each outcome by comparing the binary variables between reviewers. Depending on the inter-rater criteria, values between 0.61 and 0.8 are interpreted as substantial inter-rater reliability, and values between 0.81 and 1 are seen as almost perfect inter-rater reliability. The most relevant outcome is the slot offer. The percent agreement between the ratings for offers is 0.95 and the most restrictive measures of inter-rater reliability, such as the Cohen's Kappa or the Krippendorff's Alpha, are around 0.68. Hence, we interpret the rating for the slot offers as very reliable. Furthermore, the rating for the waiting list is very reliable, with values of 0.89 percent agreement and around 0.67 for Cohen's Kappa and Krippendorff's Alpha. The inter-rater reliability for the other outcomes is somewhat lower due to the more subjective nature of the outcomes. Still, with values for the percent agreement between 0.65 and 0.8, the reliability of the rating of the other outcomes is also substantial.

Appendix G. Natural Language Processing Outcome Classification

To classify email responses, we used the supervised Bidirectional Encoder Representations from Transformers (BERT) for sequence classification model proposed by Devlin et al. (2018) and adapted to German by Chan et al. (2020). To pre-train the model, we used 1,000 of our pre-classified emails. We chose observations for the training randomly and balanced on outcomes, such that the algorithm cannot take the likelihood of an outcome in the final sample into account. Also, we train the model on a set of anonymous emails, which do not contain the names of parents, child care centers, and center managers. The test data ($N = 100$) of our model is randomly chosen from all other remaining observations. To decide on the model parameters and to avoid over-fitting, we select the model epoch for which training loss is equal (or slightly higher) to test loss. We therefore use a total of 1,100 observations to calibrate the model and classify the remaining 11,447 observations into outcome categories.

Natural language processing algorithms are not yet capable of reliably detecting emotions in texts (such as the level of encouragement), which are often communicated between the lines or require a deep understanding of context. Therefore, we are only able to classify slot offers and waiting list offers computationally, as both are explicitly mentioned in the email texts. Results from the BERT NLP algorithm classification are presented in Table B7.

References

Bundesamt für Migration (2019). *Migrationsbericht der Bundesregierung - Migrationsbericht 2019*. Bundesminsterium des Inneren.

Bundeswahlleiter, D. (2021). Ergebnisse der repräsentativen Wahlstatistik. Statistiken zur Bundestagswahl, Der Bundeswahlleiter, Wiesbaden.

Chan, B., S. Schweter, and T. Möller (2020). German's Next Language Model. *arXiv*, preprint:2010.10906.

Clarke, D., J. P. Romano, and M. Wolf (2020). The Romano–Wolf Multiple-Hypothesis Correction in Stata. *The Stata Journal 20*(4), 812–843.

Cui, J., L. Natzke, and S. Grady (2021). Early Childhood Program Participation: 2019. First Look. Technical report, National Center for Education Statistics at IES, Washington DC.

Del Boca, D., C. Pronzato, and G. Sorrenti (2016). When Rationing Plays a Role: Selection Criteria in the Italian Early Childcare System. *CESifo Economic Studies 62*(4), 752–775.

Destatis (2021). *Bevölkerung und Erwerbstätigkeit: Bevölkerung mit Migrationshintergrund*. German Federal Statistical Office, Wiesbaden.

Devlin, J., M.-W. Chang, K. Lee, and K. Toutanova (2018). BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding. *arXiv*, preprint:1810.04805.

Dijkstra, L., H. Poelman, and P. Veneri (2019). The EU-OECD Definition of a Functional Urban Area. Technical Report, EU-OECD.

Eurydice (2019). *Key Data on Early Childhood Education and Care in Europe*. Luxembourg: Publications Office of the European Union.

Expat (2022). How to get a place in a crèche for your baby. `https://expat-in-france.com/place-creche/`. Accessed on March 27, 2023.

Felfe, C. and R. Lalive (2018). Does Early Child Care Affect Children's Development? *Journal of Public Economics 159*, 33–53.

Harvey, S. (2022). Daycare and preschool in Spain. `https://www.expatica.com/es/living/family/preschool-spain-107654/`. Accessed on March 27, 2023.

Heß, S. (2017). Randomization Inference with Stata: A Guide and Software. *The Stata Journal 17*(3), 630–651.

INKAR (2021). Indikatoren und Karten zur Raum- und Stadtentwicklung. Statistiken zur räumlichen Zusammensetzung, Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR) im Bundesamt für Bauwesen und Raumordnung (BBR), Bonn.

Jessen, J., S. Schmitz, and S. Waights (2020). Understanding Day Care Enrolment Gaps. *Journal of Public Economics 190*, 104252.

Kaas, L. and C. Manger (2012). Ethnic Discrimination in Germany's Labour Market: A Field Experiment. *German Economic Review 13*(1), 1–20.

List, J. A., A. M. Shaikh, and Y. Xu (2019). Multiple Hypothesis Testing in Experimental Economics. *Experimental Economics 22*(4), 773–793.

Malik, R., K. Hamm, L. Schochet, C. Novoa, and S. Workman (2018). America's Child Care Deserts in 2018. `https://www.americanprogress.org/issues/early-childhood/reports/2018/12/06/461643/americas-child-care-deserts-2018/`. Accessed on March 27, 2023.

NYC Department of Education (2023). Pre-K. `https://www.schools.nyc.gov/enrollment/enroll-grade-by-grade/pre-k`. Accessed on March 27, 2023.

OECD (2018). Settling In 2018. Technical report, OECD, Paris.

OECD (2020). Is Childcare Affordable? Policy Brief on Employment, Labour and Social Affairs. Technical Report, OECD, Paris.

Renfrewshire Council (2023). Apply for a free nursery or childcare place. Accessed on March 27, 2023.

Romano, J. P. and M. Wolf (2005). Stepwise Multiple Testing as Formalized Data Snooping. *Econometrica 73*(4), 1237–1282.

Romano, J. P. and M. Wolf (2016). Efficient Computation of Adjusted p-Values for Resampling-Based Stepdown Multiple Testing. *Statistics & Probability Letters 113*, 38–40.

Westfall, P. H. and S. S. Young (1993). *Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment*, Volume 279. John Wiley & Sons.