DEPARTMENT OF ECONOMICS
JOHANNES KEPLER UNIVERSITY OF LINZ

Testing for Ethnic Discrimination in Outpatient Health Care: Evidence from a Field Experiment in Germany

by

Martin HALLA

Christopher KAH

Rupert SAUSGRUBER

Working Paper No. 2115
This Version: February 2022
First Version: November 2021

Johannes Kepler University of Linz
Department of Economics
Altenberger Strasse 69
A-4040 Linz - Auhof, Austria
www.econ.jku.at

martin.halla@jku.at

Testing for Ethnic Discrimination in Outpatient Health Care: Evidence from a Field Experiment in Germany*

Martin Halla[a,b,c,d], Christopher Kah[e], Rupert Sausgruber[f]

[a] *Johannes Kepler University Linz*
[b] *Christian Doppler Laboratory Aging, Health, and the Labor Market*
[c] *Institute for the Study of Labor (IZA), Bonn*
[d] *Austrian Public Health Institute (GÖG), Vienna*
[e] *Mercedes-Benz AG*
[f] *Vienna University of Economics and Business*

February 14, 2022

Abstract

To test for ethnic discrimination in access to outpatient health care services, we carry out an email-correspondence study in Germany. We approach 3,224 physician offices in the 79 largest cities in Germany with fictitious appointment requests and randomized patients' characteristics. We find that patients' ethnicity, as signaled by distinct Turkish versus German names, does not affect whether they receive an appointment or wait time. In contrast, patients with private insurance are 31 percent more likely to receive an appointment. Holding a private insurance also increases the likelihood of receiving a response and reduces the wait time. This suggests that physicians use leeway to prioritize privately insured patients to enhance their earnings, but they do not discriminate persons of Turkish origin based on taste. Still, their behavior creates means-based barriers for economically disadvantaged groups.

JEL Classification: I11, J15, I14, I18, H51, C93.
Keywords: Discrimination, immigrants, ethnicity, health care markets, health insurance, inequality, correspondence experiment, field experiment.

* *Corresponding author:* Rupert Sausgruber, Vienna University of Economics and Business (WU), Department of Economics, Welthandelsplatz 1, 1020 Vienna, Austria; email: rupert.sausgruber@wu.ac.at.
For helpful discussions and comments, we would like to thank seminar participants at the JKUecon Staff Seminar 2019, the Austrian Labor Economics Workshop 2019, the Berlin Applied Micro Seminar, the Departmental Seminar of the University of Regensburg, and the Departmental Webinar at Boğaziçi University. The usual disclaimer applies. This project has been approved by a WU ethics panel under approval number WU-HSRP-2018-001. We gratefully acknowledge financial support from the Austrian Science Fund FWF (SFB F63 "Credence Goods"). We would like to thank Steffen Greßbach for consulting us on the process of remuneration for dentists in Germany.

1 Introduction

Members of minority groups face many challenges to their health. Relatively little is known about discrimination within modern health care markets, and whether this contributes to health outcomes.[1] We ask the question of whether physicians treat members of an ethnic minority differently. Extant evidence from a range of other settings shows a different treatment of minorities due to taste-based or statistical discrimination.[2] These two motives are notoriously difficult to separate, and in in many studies, it remains unclear whether market participants care about ethnicity *per se*, or whether it is just a proxy for something else (such as productivity). Moreover, discriminatory intentions may not be sufficient to shift outcomes, and whether discrimination exists can vary widely depending on the inner workings of a market.

The lack of systematic research on ethnic discrimination in modern health care markets is a matter of concern. Discrimination is not only unethical and illegal, but actual and perceived discrimination are also often seen as a cause of disparities in overall life satisfaction and health status (Deaton, 2008; Institute of Medicine (IOM), 2003; Pascoe and Richman, 2009). The study of ethnic discrimination is particularly warranted for publicly funded health care. This institution explicitly aims to eliminate discrimination on the basis of socio-economic status, and to protect public health by ensuring equitable, needs-based access.[3]

We report the results of an email-correspondence study in the German market for outpatient health care services. We have approached 3,224 physician offices of various specialist fields in the 79 largest cities to request an appointment. We use fictitious appointment requests with randomized patients' characteristics. Our focus is on ethnicity, as signaled by distinct Turkish versus German names, and the type of health insurance. Our main outcome variables are whether an appointment is offered, and the wait time. These allows us to infer on the *access* to outpatient care. We focus on Turks, since they are the largest ethnic minority group in Germany. They are among the most negatively stereotyped migration groups, and there is compelling evidence from different contexts

[1] Historically, racism is well documented in medicine. One of the best studied cases is the United States and its black-white disparities in health (Byrd and Clayton, 2001). For instance, Alsan and Wanamaker (2017) document how the so-called *Tuskegee Study of Untreated Syphilis in the Negro Male* (conducted between 1932 to 1972) has caused medical mistrust among older black men, and contributed to a reduction in their life expectancy.

[2] Evidence for ethnic discrimination has been observed, among others, in hiring decisions (Riach and Rich, 2002), behavior at the workplace (Glover *et al.*, 2017; Hjort, 2014), the education sector (Milkman *et al.*, 2015; Boisjoly *et al.*, 2006; Farkas, 2003), the housing market (Auspurg *et al.*, 2019; Ewens *et al.*, 2014), product markets (Doleac and Stein, 2013; Zussman, 2013), the sharing economy (Edelman *et al.*, 2017), financial services (Stefan *et al.*, 2018), bureaucratic behavior (Hemker and Rink, 2017; Giulietti *et al.*, 2019), and discretionary favors in the marketplace (Mujcic and Frijters, 2021). For surveys on hiring discrimination see Baert (2017); Neumark (2018).

[3] See, for instance, Fleurbaey and Schokkaert (2011); Moscelli *et al.* (2018); Johar *et al.* (2013); Wagstaff and Van Doorslaer (2000).

that they are exposed to discrimination (Antidiskriminierungsstelle des Bundes, 2017).

The German health care system ensures universal coverage of the population for health care services in a two-tier insurance system. Close to 90% of the population is covered by mandatory insurance in the statutory health insurance tier (SHI). The remainder of the population has private insurance (PHI). Access to PHI is limited to those who are not compulsorily insured in the SHI tier, who must either prove sufficient income to afford private insurance, or belong to an occupational group that is not covered by SHI.[4] There are major differences across, but not within, insurance tiers. Most importantly, physician receive a higher remuneration for providing the same service to a patient with PHI (as compared to one with SHI). Estimates suggest that provider fees for comparable services differ between the tiers by a factor of two and more.

Statistical or third-degree price discrimination obtains when profit-maximizing firms under imperfect information sort individuals on observable group averages, such as ethnicity, as a signal for profit opportunities. The German institutional setting provides a natural testing ground for ethnic discrimination in an outpatient health care market. Three features facilitate a disentanglement of statistical and taste-based discrimination. First, conditional on insurance, physician have full information about expected profits and face no payment risk. In the SHI system, the benefits catalog for patients and the fee schedule for physicians are fully harmonized through negotiations between the associations of health care providers and the health insurers. As a result, all physicians providing service to SHI patients are reimbursed according to the same uniform scheme. In the PHI system, patient benefits and fees are also highly harmonized, although physicians have somewhat greater leeway. Second, there is no insurance-based sorting on the supply side of the market. The vast majority of physicians accept both publicly and privately insured patients.[5] Third, primary care physicians traditionally do not have a gatekeeper function. Patients can freely choose and directly access both primary and secondary care providers. This allows us to focus on a one-to-one contact between a patient and a physician without the involvement of a third party.

Our experimental setup guarantees that there is no uncertainty about a patient's insurance status. Every patient is either privately or publicly insured, and we have full control in our research design. This setup leaves little room for statistical discrimination and provides a valuable opportunity to examine ethnic discrimination based on taste. By restricting contact to a pre-determined email protocol that is constant across all treatment cells, we rule out the possibility that prior beliefs of the experimenters may bias the results

[4]Historically, SHI has been introduced by the chancellor of the German Empire, Otto von Bismarck, in 1884, to gain the votes from blue-collar workers (Bauernschuster *et al.*, 2020). Other occupational groups such as civil servants and self-employed were not eligible for SHI. Busse *et al.* (2017); Blümel and Busse (2020) provide a thorough description of the German health care system.

[5]There is only a negligible market segment servicing only PHI patients. We have not sampled from this segment for our study. The reasons for the lack of strong market segmentation could be market access regulation, but also that there are relatively few patients in PHI.

through some type of uncontrolled communication. Finally, our study has high statistical power, and we are careful to measure ethnic discrimination as cleanly as possible.[6]

To get a broad picture, we include several groups of specialists in our study. We contact dentists, ophthalmologists, dermatologists, and orthopedists. For one, these groups differ with respect to their respective opportunities to "sell" additional services beyond what is requested in the email, and also in terms of competition. We use common and mostly harmless symptoms like a loose filling in the email to the dentist, or an inflammation of the eye in the message to the ophthalmologists. The symptoms are deliberately chosen so that they are likely to require only a single treatment. In this way, we constrain beliefs about continuing appointments as a possible unobservable cause of statistical discrimination. The emails are written concisely in flawless German to avoid another layer of potentially confounding influence from language.[7] We use the following wording (with different text for each treatment/type of physician in parentheses): "*Dear Madams and Sirs, I need an appointment urgently, please, because I have [symptom]. I'm on currently on vacation and will return home on Monday. I am insured with [insurance]. When would be the next possible appointment? Kind regards, [Name]*". Because the research question is touching on a sensitive topic, we are careful so that the experiment is kept from being noticed. We sample physicians only from cities with more than 100,000 inhabitants to guarantee a high level of anonymity of the study.

Our main results are as follows. First, we observe that the manipulation of perceived ethnicity has no effect on appointment rates, wait times, or any other outcome. This result holds independently of the insurance type and the specialists' group. Second, the type of insurance looms large. SHI holders get an appointment in 41% of the cases. For PHI counterparts the rate is 32% higher and amounts to 54 percent. Further differences between SHI and PHI patients are in regard to the likelihood of receiving a response (72 vs. 75%), the length of time they have to wait for an appointment (plus a third of a day), and how often they are told to expect a long wait in the physician's office (22 vs. 16%). Apparently physicians do have the ability to prioritize treatment, and they make extensive use of it. The most plausible mechanism are the higher provider fees in PHI, which make it beneficial to treat private patients better. Importantly, however, the ethnicity of the patient plays no role in this.

Our results are robust across important sub-samples. To look at effects of the general level of xenophobia as a likely amplifier for discriminatory intentions, we consider heterogeneity in the regional vote share of the largest right-wing populist party. Among others, we also consider sample splits by city size, the share of foreigners, the share of Turks, and the physicians' type. In neither sub-sample we find evidence of ethnic discrimination.

[6]With implemented sample size, we have ex-ante calculated to reject the null hypothesis of ethnic non-discrimination at a minimal detectable effect size of 0.044 ($\alpha = 5\%$ and $1 - \beta = 80\%$).

[7]For a discussion of statistical discrimination as signaled by language, see Balsa and McGuire (2001).

This result may seem surprising at first, but it is plausible on closer inspection. Our favored explanation is that in the case of universal coverage and patients free choice of physician, transparency of insurance deprives the market of any grounds for statistical discrimination. In this setting, price regulation cum quality competition prevent any attempts to discriminate based on taste. Because prices are regulated, physicians compete for market share on non-price dimensions such as effort, quality of treatment, and waiting time (McGuire, 2000).

To substantiate this intuition, we exploit local variation in specialist-specific market densities to measure competitive pressure. Theory predicts that quality is increasing in the intensity of competition and in the regulated price (see, for example, Gaynor (2007)). In line with this prediction, we do observe ethnic discrimination among SHI holders in low-competitive segments of the market. There is also some evidence that competition imposes a limit on the extent of insurance-based discrimination, i.e., patients with private insurance are less favored in a more competitive environment. While the analysis accounts for unobserved heterogeneity at the city-level, we cannot assume our measures of competition to be fully exogenous, and we must interpret these results with care. Yet, the results suggest that overt discrimination does not exist and unlikely gives rise to health disparities between the Turkish minority and the resident population. In the spirit of Phelps (1972) and Arrow (1973), this is a consequence of universal health care, leaving little room for statistical discrimination. Following the intuition of Becker (1957), however, the result also reflects the fact that physicians compete for patients.

Our study makes several contributions. First, we add to the small literature testing experimentally for ethnic discrimination in the access to outpatient health care services. Existing evidence comes from three telephone studies in the US providing mixed evidence. Most closely related to our study, Sharma *et al.* (2015) conduct simulated patient calls with imposed variation to insurance status, race, and sex to a random sample of 922 primary care physicians to assess appointment availability. The authors do no find robust differences across ethnic groups.[8] Our study improves on these existing field experiments. First, we have a much larger sample size than these prior studies. Second, we use the internally more valid appointment request mode via email. Third, we create a specific scenario in the German health care setting, which facilitates a clean isolation of taste-based discrimination. Fourth, we explore treatment effect heterogeneity across areas,

[8]Two other US studies use a similar telephone-based approach, but in comparable smaller and less representative settings. First, Kugelmass (2016) contacts 320 New York based psychotherapists via voice messages. The author uses racially distinctive names, as well as race- and class-based speech patterns to manipulate race, social class, and gender. Holding insurance status constant, the author finds that black patients are less likely (as compared to white patients) to be offered an appointment. Wisniewski and Walker (2020) contact 804 primary care offices in two urban centers in Texas. In their study, they vary race but do not hold insurance status constant. Their results show that black and hispanic patients are more likely to get an appointment, but that these groups have to wait longer for an appointment, and are asked more often about their insurance.

which differ in terms of important dimensions such as the overall level of xenophobia and competition in the specific outpatient health care market.

Our finding of no taste-based discrimination against people with a Turkish migration background is in stark contrast to results from corresponding testing studies in other contexts, and to survey based evidence on health. The former demonstrate discrimination of Turks in Germany on the labor market (Kaas and Manger, 2011; Weichselbaumer, 2020), the housing markets (Auspurg *et al.*, 2019), the carpooling market (Carol *et al.*, 2019), and in the communication with welfare offices (Hemker and Rink, 2017). An example for the latter is Igel *et al.* (2010), who show based on data from the *German Socio-Economic Panel* that migrants with Turkish background report significantly poorer health, and perceive themselves more discriminated than any other minority groups. We differ from this type of survey study in terms of method. Survey responses may generate biased results if they reflect inaccurate beliefs.[9] Our approach avoids this concern and also facilitates an interpretation holding insurance status constant.

A second contribution of our study is to quantify the causal impact of insurance status on the access to outpatient health care services. Whether there is first- and second-class medicine is a hotly debated topic in German politics. Not surprisingly, there has been work to measure differences between the insurance tiers. Our results confirm and expand existing evidence. Most closely related is the field experiment by Werbeck *et al.* (2021), in which fictitious patients with SHI versus PHI contacted 991 private practices in 36 German counties to schedule appointments for allergy tests, hearing tests and gastroscopies. In line with our results, the authors find that patients with PHI are more likely to be offered an appointment (plus 4%), and have to wait less (minus 100%). Quantitatively, Werbeck *et al.* (2021) find (as compared to our results) smaller effects of PHI on appointment rates, but a stronger effect on wait times conditional on being offered an appointment. A potential explanation for this differential finding is the specific choice of specialists and requested treatments. Similar studies with comparable fewer observations ($N < 200$) have been conducted by Heinrich *et al.* (2018) and Lungen *et al.* (2008). There is also comparable experimental evidence for the German inpatient sector (Kuchinke *et al.*, 2009; Schwierz *et al.*, 2011). None of these studies account for ethnic discrimination, as we do.[10] This previous evidence, together with our finding of non-ethnic discrimination, suggests that German physicians favor privately insured patients not because of unobserved characteristics, but due to the higher compensation they receive.

[9] Bohren *et al.* (2019) provides a careful discussion of the causes and consequences of systematically incorrect beliefs in the context of discrimination.

[10] There is also evidence on the importance of insurance status from experimental audit and correspondence studies in the US (see, for instance, Bisgaier and Rhodes, 2011; Polsky *et al.*, 2015). Some studies evaluate "cream skimming" based on other socioeconomic characteristics. For instance, Angerer *et al.* (2019) show that Austrian physicians positively discriminate patients with a university degree. This observation is consistent with statistical discrimination, because a university degree can signal higher willingness to pay for services that are not covered by insurance.

Finally, our findings are also relevant from an ethical and legal perspective. Differential treatment based on insurance status is unethical, among others, because physicians are required by their professional code of conduct to treat all patients equally. In Germany, as in many other countries, it is also illegal because any discrimination, including preferential treatment of PHI patients, is prohibited under civil law as well as under the *General Equal Treatment Act*.[11] The presence of discrimination against such provisions should caution regulators to be mindful of economic incentives.[12]

The paper proceeds as follows. Section 2 provides a brief discussion of the relevant aspects of the German health care system. Section 3 describes the correspondence testing study, including the sampling procedure, the experimental design, and provides descriptive statistics on covariates and outcome variables. Section 4 presents our estimated treatment effects. In this section, we first discuss average treatment effects and then test for treatment heterogeneity across different sub-samples. Section 5 concludes the paper.

2 The German health care system

In Germany, health insurance is mandatory. Employees are enrolled in the statutory health insurance (SHI, *Gesetzliche Krankenversicherung*) by default. Self-employed persons, civil servants and persons whose yearly income exceeds a certain threshold of currently €62,550 can opt for SHI, or must take out private health insurance (PHI, *Private Krankenversicherung*). In 2019, 10.5% of the total German population had full private insurance (Verband der Privaten Krankenversicherung, 2020). A small percentage of the population of about one percent (members of the military, police, and refugees) is insured through special government schemes.

Individuals insured in SHI are members of one of over one hundred non-profit insurance funds (*Krankenkassen*) that contract with providers. They can freely choose between insurance funds and between SHI-accredited health care providers. In accordance with the non-discriminatory solidarity principle (*Solidaritätsprinzip*), the system is financed by a flat earmarked payroll tax of 14.6%, while benefits are based on need. Consequently,

[11]Formally, the provisions of the *General Equal Treatment Act* (*Allgemeines Gleichbehandlungsgesetz*, AGG §2 Abs. 1(5)) extend to contracts between physicians and (statuary and private) health insurance patients. Anti-discrimination in access to health care also follows from Articles 21 and 35 of the *Charter of Fundamental Rights of the European Union*. Article 21 recognizes the right to be free from discrimination, including on the grounds of sex, racial or ethnic origin, and religion or belief. Article 35 guarantees the right of access to healthcare under the conditions established by national laws and practices. See, Antidiskriminierungsstelle des Bundes (2017).

[12]A related but distinct strand of literature examines waiting times for inpatient treatments, both theoretically (see, for instance Iversen, 1993; Brekke *et al.*, 2008) and empirically (see, for instance, Cullis *et al.*, 2000; Siciliani *et al.*, 2014; Monstad *et al.*, 2014). This setting gives rise to different aspects, since in most systems hospitals are (at least partly) financed based on diagnosis-related groups (DRG) fee-per-case systems. Whether or not discrimination is costly to a provider heavily depends on the specific regulations.

high and low risks are in the same pool.[13] In SHI, the provider fees are fully harmonized according to a uniform fee schedule called "Unified Assessment Scale" (*Einheitlicher Bewertungsmaßstab*, EBM). The patients receive benefits in kind, and the physicians charge the insurance funds based on the EBM. Physicians on the one hand and insurance funds on the other are organized in associations that negotiate the fee schedule centrally. SHI-physicians are by law members of their respective regional association. They are not allowed to charge patients directly for services from the SHI benefits catalog.

In PHI, coverage and premiums differ by insurance plan, but provider fees are comprehensively regulated, too. The so-called "Medical Fee Schedule" (*Gebührenordnung für Ärzte*, GOÄ, and the *Gebührenordnung für Zahnärzte*, GOZ, for dentists, respectively) is a legal directive issued by the federal government and centrally negotiated between the *German Medical Association* (*Bundesärztekammer*) and the Private Health Insurance Association. The GOÄ/GOZ establish base prices for individual services. Physicians can add markups based on the complexity of the treatment. The markup must not exceed the base price by factor 3.5, and there is a threshold factor of 2.3, the exceeding of which requires justification by the specifics of the medical treatment. Such treatments can only be carried out with the patient's consent and are sometimes not covered by insurance plans. These provisions imply that the fees for each service are virtually unified at the base price $\times$ 2.3. In fact, the vast majority of cases bunches at this threshold (see, Walendzik *et al.* (2009)). In PHI, providers bill patients directly, who then seek reimbursement from their insurance company. PHI plans cover a range of treatments that go beyond the benefits catalog of SHI. Examples are comprehensive coverage for fillings and dentures, increased preventive health examinations, or amenities and guaranteed treatment by a chief physician in a hospital.

The fee schedules of SHI and PHI are not easily comparable, though. A major difference is that GOÄ/GOZ follow a fee-for-service principle, while the EBM in the statutory regime is based on diagnosis-related groups. Evidence suggests that physicians earn different fees for the same service. Walendzik *et al.* (2009) find that, on average, physicians receive a 2.3 times higher fee for the same medical services for treating PHI than SHI

[13]Further features of SHI are that non-working family members are covered free of charge, the statutory tax burden is divided equally between employers and employees, and earnings above €58,050 per year are exempt from paying contributions. The contributions accrue to a central pool (called *Gesundheitsfonds*), which reallocates the money to the insurance funds to cover expenditure.

patients.[14] To some extent, physicians may also be able to offer additional services, such as dental cleanings at dentists, especially if this is included in the insurance plan. Some physicians say quite openly that treating only SHI patients would not be profitable and that private patients would cross-subsidize SHI patients. This suggests that there are incentives to prioritize privately insured patients. There is the familiar tale of separate waiting rooms in a physician's office, with shorter waiting times for privately insured patients. However, the evidence is mixed at best, and research designs are often not strong enough to causally identify the impact of insurance (see Werbeck *et al.* (2021) for a discussion on this point).

In summary, the SHI system eliminates essentially all grounds for statistical discrimination based on ethnic origin or other personal characteristics. Also, within PHI, there are only minor possible causes for this source of discrimination. For example, one would have to assume that physicians' beliefs about the prospect of selling additional treatments (or of overtreatment) differ among privately insured patients of different ethnic backgrounds.

3 Correspondence testing study

The data we use in this paper come from an experimental study we conducted in May 2017 and February 2018. In total we contacted 3,224 physicians from the 79 largest cities in Germany (see Appendix Section A.1 for the full list of cities).

3.1 Sampling procedure

We have sampled physicians located in German cities with a population in excess of 100,000. The rationale for this restriction is that we want to minimize the risk that, in small towns with a low number of medical specialists, physicians know each other and find out in direct personal conversations that they are part of our corresponding testing study. We focus on four types of physicians: dentists, dermatologists, ophthalmologists, and orthopedists. For each city and type, we draw a random sample from the "Gelbe

[14] To arrive at this number, Walendzik *et al.* (2009) use a provision that enables SHI-insured persons to opt for cash benefits instead of benefits in kind in order to claim insurance benefits outside the SHI system. They calculate markup factors between PHI and SHI of 1.52, 1.69, 1.73 and 1.92 for ophthalmologists, dermatologists, otorhinolaryngologist and orthopedists, respectively. Factors are even higher for urologists (2.10), gynecologists (2.13), psychiatrists (2.29), internal medicine specialists (2.40), pediatricians (2.45), general practitioners (2.64) and radiologists (2.99). However, as noted by the authors, the results must be interpreted with care because they do not account for possible selection effects at the patient's side. Based on private communication with a dentist in Frankfurt am Main, for the treatment of the loose filling the provider fees amounts to about €127 and €66 for PHI and SHI patients, respectively. Own computations for the other medical treatments in our study confirm a factor of about 2, which is in line with the figures by Walendzik *et al.* (2009).

Seiten."[15]

The target sample size for each specialization per city is proportional to the universe.[16] We have aimed for a random sample of 30%. If the number of available physicians in a given group has fallen short of this target ratio, we have sampled the maximum available number. By definition, larger cities contribute a higher number of physicians. For instance, 9.5% of all sampled physicians are from Berlin, 5.0% from Munich, and 4.9% from Hamburg. The median share per city is 0.96%. This procedure yields a total sample of 3,224 physicians, which comprises 1,552 dentists, 405 dermatologists, 638 ophthalmologists, and 629 orthopedists.

We do not aim at individually identifying discriminating physician offices, and we send only one request to each physician. This procedure is no threat to our identification, since physicians supply service to both types of insurance. While this implies that we can only observe the average discrimination in the market, we thereby avoid the important concern that physician offices could become wary of similar looking emails (see, Kessler *et al.* (2019)).

3.2 Covariates

The information provided on the "Gelbe Seiten" is rather sparse. It contains the name of the physician's office, contact information, in most of the cases opening hours, and the rank on the "Gelbe Seiten." As covariates we use a binary indicator for group practice (derived from the office's name or information on website), the total weekly office hours, and the physician offices rank.[17] We further use physicians' type and city fixed-effects.

3.3 Experimental design

We contacted physician offices by email and asked for an appointment. The senders are four different fictitious patients. In one dimension, we vary the insurance status (SHI vs. PHI), while in the other dimension we vary the perceived ethnicity (German vs. Turkish) via the name of the sender.[18] As statutory insurance we chose "AOK", the most common in Germany. As private insurance, we picked "Allianz AktiMed Best". This is a premium PHI plan that provides universal coverage including the refund of relatively

[15]This website (`https://www.gelbeseiten.de`) is the German equivalent of "Yellow Pages". Registration at this website is free of charge for any firm, including physician offices. A paid premium account allows uploading photos and also improves the rank.

[16]The total number of physicians (universe) is computed based on data from the *Bertelsmann Stiftung* (hyperlink, last accessed on).

[17]In about 10% of cases, information on office hours is lacking (i.e, neither available on "Gelbe Seiten" nor on the office's website). In these cases, we have imputed this information with the sample mean.

[18]Turks in Germany are the largest ethnic minority group in Germany, and also the largest Turkish community in the Turkish diaspora. Germany is a country of 83 million inhabitants and hosts about 1.48 million Turkish nationals. The number of people with migrant background from Turkey is about twice that number (Statistisches Bundesamt (Destatis), 2019; Schührer, 2018).

large markup factors. We choose "Christian Schmidt" as German and "Ahmet Yilmaz" as Turkish name of senders. "Schmidt" is the most common surname in Germany, while "Christian" ranked top among male given names in 1973. "Yilmaz" is the most common Turkish surname in Germany, while "Ahmet" is among the most popular given Turkish names.[19]

Contact via email might not be the most common practice; however, it circumvents concerns that arise with telephone studies. Bertrand and Duflo (2017) point out that when test persons know the purpose of a study "*[...] this may generate conscious or subconscious motives among auditors to generate data consistent or inconsistent with their beliefs about race or gender issues*". Emails are a method to reliably identify the effects of varying our parameters of interest. In the emails, we indicate that the sender is currently on vacation and needs an appointment as soon as possible upon return home. The fact that the appointment request is coming from vacation makes the email contact (instead of the usual phone call) more credible. Our overall response rate of 0.74 shows that email are an accepted way to ask for an appointment.

The symptoms we use in the emails vary by specialist type. They are quite harmless, but require relatively prompt treatment. In the email to the dentist, we request an appointment due to a loose filling. The other symptoms are an inflammation of the eye (ophthalmologist), back pain (orthopedist), and a rash in genital area (dermatologist). The exact wording of the emails can be seen in the Appendix Section A.2. The symptoms are deliberately chosen such that a regular patient would neither directly attend emergency consultation hours nor need a referral from a general practitioner to a specialist.

Each physician in our sample was randomly assigned to only one of four fictitious senders. Table 1 lists our covariates by treatment and demonstrates the balanced distribution across treatments. The emails were sent in two waves in May 2017 and February 2018. Each mailout included all treatments in equal shares. To minimize physicians' cost, we immediately cancelled an appointment upon confirmation. Hence, we can rule out the possibility that other patients were denied an appointment because of our study. Our study is also minimally invasive in that dealing with an appointment request is a standard routine that takes only a few minutes of the scheduler's time in a physician's office.[20]

3.4 Outcome variables

Our outcome variables are based on the physicians' response emails or lack thereof. We have encoded the receipt of an answer. Conditional on response, we measure the days to response. The most important outcome variable is whether or not an appointment is

[19] Emails were sent from `christian.m.schmidt1973@gmail.com` and `ahmet.m.yilmaz1973@gmail.com`, respectively.

[20] We encode the data in an anonymous fashion and the estimation results in Section 4 refer to aggregate data.

offered. Conditional on an appointment, we calculate the duration between request and appointment, and any information on the expected waiting time in the physician's office. Finally, we also track whether physicians answer the email, even if they do not offer an appointment.

Table 2 summarizes the average outcomes by type of physician.[21] Between 0.70 and 0.75% of physicians send a response. The response time is fast (usually within a day). These figures confirm that email queries are well suited for our research design. The share of physicians offering an appointment is substantially lower and varies across types. Dentists are the most likely to offer an appointment (0.53), and orthopedist the least likely (0.40). A significant share of responses explicitly mentions that patients have to expect a wait in the physician's office. There is substantial variation in this outcome across types. The "long-wait" appointments are most common among ophthalmologists (0.52) and least common among dentists (0.07). The time to an appointment is rather low. On average, physicians offer to see a patient after six days. Finally, conditional on declining an appointment, every second physician at least answers the email. This share hardly differs across types.

4 Estimation results

We first discuss average treatment effects. Thereafter, we check for potential treatment effect heterogeneity by replicating our estimation analysis in several sub-samples. Most importantly, we distinguish between cities with a low versus high competition in the market for a given medical speciality.

4.1 Average treatment effects

Figure 1 shows unconditional means for all outcomes across fictitious patients.[22] Table 3 summarizes estimation results from OLS regressions with and without covariates for five outcome variables, respectively. Listed coefficients for binary outcome variables have to be interpreted as percentage points differences relative to Ahmet Yilmaz with statutory insurance (base group). Coefficient for non-binary variables are unit-changes relative to base group.

Panel (a) of Figure 1 shows that there is some variation in the likelihood of receiving an answer across fictitious patients. Response rates are higher in PHI than SHI, both, for Christian Schmidt and Ahmet Yilmaz. The estimated effect is about plus 5 percentage points (see column 1 and 2 of Table 3). Conditional on insurance type, there is no significant differences between the response probability to Christian Schmidt and

[21] Alternatively, Appendix Figure A.1 provides an equivalent graphical depiction.

[22] For our two non-binary outcomes, we plot in Appendix Figure A.3 estimated kernel density functions.

Ahmet Yilmaz. Panel (b) shows that all senders who receive a response are on average informed within a day. Notably, Christian Schmidt with PHI receives a reply only about 8 hours later (see column 3 and 4 of Table 3). We interpret this finding as indicating that physicians are making more effort to schedule an appointment for patients with PHI.

Most importantly, in Panel (c) we see that senders with private insurance are more likely to receive an appointment, irrespective of their name. Both, Christian Schmidt and Ahmet Yilmaz with private insurance receive an appointment in about 54% of the cases. For those with statutory insurance, these figures are 42 and 40% for Schmidt and Yilamz, respectively. Thus, holding PHI increases the likelihood of receiving an appointment by 12 to 14 percentage points. This is a sizable discrimination of patients with SHI of about 31%. These results are robust to the inclusion of control variables estimates (see columns 5 and 6 of Table 3).[23] Appendix Table A.1 lists more detailed estimation output for four specifications, with increasing stepwise the number of covariates.

Conditional on receiving an appointment, privately insured patients also tend to receive "better" appointments. First, the duration between request and appointment is somewhat shorter. For Christian Schmidt the reduction is almost half of a day (or 6.5%), which is significant at the 10% level. For Ahmet Yilmaz the reduction is comparably smaller (a fifth of a day or 3.4%) and statistically insignificant. Second, the expected wait in the physician's office is shorter for private patients. They are 6.3 (Schmidt) and 7.3 (Yilmaz) percentage points less likely to receive a response stating that they can expect a long waitime than those with SHI. Again, conditional on insurance type, there are no statistically significant differences between Christian Schmidt and Ahmet Yilmaz.

Im summary, there is no evidence for ethnic discrimination by German physicians in patient appointment scheduling. This holds true even after accounting for control variables.[24]

4.2 Heterogeneity in discrimination

We now test for treatment effect heterogeneity by splitting our sample along different dimensions. We consider the regional vote share of the largest right-wing populist party, city size, the share of foreigners, the share of Turks, former East versus West Germany, and the physicians' type. Given the evidence on the impact of physician-patient racial concordance on health outcomes (Alsan *et al.*, 2019), it would be also interesting to

[23]These results are plausible in the light of previous findings. For example, Polsky *et al.* (2015) study the effect of an increase in Medicaid reimbursement. They find a sizeable effect on appointment rates (10% increase in reimbursement rate leads to 1.25 percentage point difference in appointment rate). Assuming that provider fees in PHI are approximately twice as high as in SHI, i.e. 100% increase, this would translate to a 12.5 percentage points increase in appointment rates.

[24]One possible concern is that our intervention may not be strong enough to sufficiently change recipients' perceptions about the ethnicity of email senders. However, this interpretation is at odds with the fact that other studies, who have employed the same method with very similar names, have measured very strong effects. For example, Hemker and Rink (2017) use "Michael Schäfer" vs. "Mustafa Yilmaz."

distinguish between Turkish and German physician.[25] We have categorized physicians in our sample based on their names as Turks (or other foreigners). Unfortunately, the number of Turkish physicians is too low (76 cases or 2.3%) for a statistical analysis.

4.2.1 General level of xenophobia and share of foreigners

It could be that, on average, ethnic discrimination is not observed, although it is present in areas with high levels of xenophobia. To explore this possibility, we split our sample by the median vote share for the *Alternative for Germany* (*Alternative für Deutschland*, AfD) in general elections in 2017. The AfD is a right-wing populist political party known for its oppositional attitude towards immigration (Cantoni *et al.*, 2019). The AfD vote share varies across the cities in our sample between 4.9% and 24.3%, with a mean of 11.2%. Panel (a) of Figure 2 summarizes estimated effects for the outcome appointment in the two subsamples with a AfD vote share below and above the median. Estimated effects have to be interpreted relative to the base group (Yilmaz, SHI). Appendix Table A.2 provides more detailed estimation output for this, and all other outcomes. We see no evidence for differences in ethnic discrimination between cities with a low versus high share of AfD votes. Also, discrimination based on insurance status is very comparable across regions.[26]

We perform an equivalent analysis for small versus large cities. Large cities are defined as cities with a population of 500,000 inhabitants or more. Metropolitan areas are known to be more open to immigration and ethnic diversity (Dustmann *et al.*, 2019). However, we do not find any evidence for significant differences in discrimination due to ethnicity (or insurance status) between small and large cities (see Panel (b) of Figure 2 and Appendix Table A.3). Finally, we distinguish between cities according to the share of foreigners (i.e., residents without German citizenship), and the share of residents with Turkish citizenship. The share of foreigners varies across the cities in our sample between 6.4 and 40.7%, with

[25] Alsan *et al.* (2019) randomize black men to black or non-black physicians. Patients who meet a racially concordant doctor have a higher demand for preventive services. The authors argue that this effect is driven by the extent to which physicians are able to correct false beliefs and not by taste-based discrimination. Chandra and Staiger (2010) study treatment disparities for heart attacks based on a different research design using observational data and also reject the role of prejudice against black patients.

[26] This assessment is confirmed by proper statistical inference based on the following estimation model,

$$
\begin{aligned}
app_i = {} & \eta^1 \cdot \text{Schmidt PHI} \times \text{AfD}^{\text{low}} + \eta^2 \cdot \text{Schmidt PHI} \times \text{AfD}^{\text{high}} + \eta^3 \cdot \text{Schmidt SHI} \times \text{AfD}^{\text{low}} \\
& + \eta^4 \cdot \text{Schmidt SHI} \times \text{AfD}^{\text{high}} + \eta^5 \cdot \text{Yilmaz PHI} \times \text{AfD}^{\text{low}} + \eta^6 \cdot \text{Yilmaz PHI} \times \text{AfD}^{\text{high}} \\
& + \eta^7 \cdot \text{Yilamz SHI} \times \text{AfD}^{\text{low}} + \eta^8 \cdot \text{Yilamz SHI} \times \text{AfD}^{\text{high}} + \sum_p \lambda_p + \sum_c \eta_c + \epsilon_i,
\end{aligned} \tag{1}
$$

where AfD^{high} and AfD^{low} are binary indicators for observations from cities with an AfD vote share below and above the sample median, respectively. λ_p and η_c represent physician's type and city fixed-effects. We test $\eta^1 = \eta^2$, $\eta^3 = \eta^4$, $\eta^5 = \eta^6$, and $\eta^7 = \eta^8$. Estimation output is available upon request.

a mean of 17.1%. The mean share of Turkish nationals is 2.7%. We find no evidence for significant differences in discrimination due to ethnicity (or insurance status) between cities with a low versus high share of foreigners (see Panel (c) of Figure 2) or a low versus high share of Turks (see Panel (d) of Figure 2).

4.2.2 Market concentration

According to Becker (1957), competition is a means of curbing discrimination. Competition in the German outpatient health care market could, on average, be strong enough to eliminate ethnic discrimination. To explore this dimension, we rely on city-level physician-type specific densities to proxy varying degrees of market concentration within our sample. Table 4 shows the average densities (i.e, number of physician per 100,000 inhabitants) across cities and by type. There is substantial variation across both dimensions. We observe the lowest density for dermatologists (5.9), and the highest for dentists (67.0). Within in each group there is also sizeable variation across cities (see minima and maxima).

Clearly, the physician density is the result of supply and demand, and we should not simply treat this variable as exogenous. To address this concern, we refer to *type-specific median* densities. This variable varies within cities, across medical specialities, and enables us to control for city fixed-effects. Thus, we allow for unobserved heterogeneity at the city-level. This captures aspects such as a general preference for a high or low level of health-care utilization. Panel (c) of Figure 3 summarizes estimated effects for the outcome appointment in the two subsamples with a market concentration below and above the median.[27] In line with the theoretical prediction, we find evidence for ethnic discrimination in markets with a low concentration. Conditional on statutory insurance, Christian Schmidt is 6.0 percentage points more likely to receive an appointment as compared Ahmet Yilmaz. This effect is marginally significant (P-value = 0.078). In contrast, in markets with a high concentration, this difference is economically and statistically insignificant. Based on a pooled estimation with a full set of interactions, we find that the estimated probability of an appointment for Ahmet Yilmaz with statutory insurance is statistically significant different between markets below and above median market concentration (P-value = 0.039).[28] For all other fictitious patients, there are no statistically significant effects across markets below and above median market concentration. In sum, we interpret this as suggestive evidence for ethnic discrimination in health care markets with low levels of competition.

Notably, the point estimates for discrimination due to insurance status are also larger in areas with a low market concentration. Both, Christian Schmidt and Ahmet Yilmaz

[27] Appendix Table A.6 provides more detailed estimation output for this, and all other outcomes.

[28] This is based on the statistical test of $\eta^7 = \eta^8$ in an equivalent estimation model as shown in footnote 26.

with private insurance have a higher predicted appointment probability in markets with a low levels of market concentration as compared to this with high levels (see Panel (c) of Figure 3). However, these differences are not statistically significant at conventional levels.

4.2.3 Types of physicians

In a final step, we explore heterogeneity in discrimination across types of physicians. Therefore, we replicate our analysis in the four respective sub-samples. Panel (d) of Figure 3 summarizes estimated effects.[29] Two main findings emerge. First, there is no evidence for ethnic discrimination in either subsample. Second, there is discrimination of patients with statutory insurance by all types with the exception of dentists. We can only speculate why dentists do not show any discriminatory behavior. As indicated by the market density figures, discrimination could be eliminated through more competition. Dentistry is the medical field withe the highest density (see Table 4) in our sample. Another explanation could be that dentists can generate more revenue than other specialists by offering additional services. However, this explanation is not particularly plausible, as such services tend to be covered by PHI, while SHI patients would have to pay for them out of their own pockets.

5 Conclusions

Our study contributes to a broader literature on the causes of minority health disparities. Existing studies suggest that access-related barriers are likely the most significant barriers to equitable health outcomes, and that disadvantaged minority groups are less likely to have a regular primary care provider (Institute of Medicine (IOM), 2003). Our study reveals that there is no taste-based discrimination of Turks in the German market for outpatient health care.

This result is perhaps surprising, as it contrasts with the findings of correspondence studies in other contexts. We do not interpret our results as indicating that members of ethnic minorities do not experience discrimination. Rather, we believe the most likely explanation is that there is little room for statistical discrimination in the German health care setting. Moreover, the breakdown of our sample by market density suggests that there is competition among physicians for patients that disciplines discrimination based on taste.

The transparency of patients' insurance status and competition among providers are two conditions that apply in the vast majority of modern healthcare markets, not just the German case. In particular, transparency reflects the situation not only in systems of

[29] Appendix Table A.7 provides more detailed estimation output for this, and all other outcomes.

universal care. In countries with multi-tier systems, including the US with parts of the population being uninsured, patients with insurance want to signal their status, and health care providers obtain information on the insurance status of patients usually by asking. Provider choice is also an important element of outpatient health care in most countries. For instance, US Medicaid beneficiaries may obtain services from any qualified provider. In managed-care plans including Medicare, choice is still guaranteed, but it is restricted to the provider network. Note also that in insurance-driven health care markets, prices per service tend to lie above marginal cost to assure participation of physicians, and rejecting a patient with eligible insurance is generally costly (McGuire, 2000). Our study suggests that because of these features, health care markets are less prone to discrimination as compared to other markets, where there is more uncertainty or insufficient competition.[30]

While our results clearly matter in regard to equal treatment, they are less clear in terms of welfare. Importantly, physicians' compensation has shown to matter for health care supply. Clemens and Gottlieb (2014) evaluate a Medicaid reform and estimate an elasticity of health care of 1.5 with respect to physicians' reimbursement rates. Similarly, Polsky *et al.* (2015) find that a 10% increase in Medicaid reimbursement rates leads to 1.25 percentage point difference in appointment rates. For Germany, Sundmacher and Ozegowski (2016) estimate that a one percentage increase in proportion of residents with PHI increases the density of specialists and general practitioners by 2.1 and 1.3%, respectively. A welfare analysis of whether a health care system should further harmonize reimbursement rates to curb unequal treatment across insurance type would need to account for both, the social welfare weights placed on ethnic-minority members and the behavioral effects of physicians' compensation.

Such structural differences in physician compensation across insurance types create means-based barriers for economically disadvantaged groups. In Germany, the average person sees a physician more than nine times a year, and 87% of the population makes use of outpatient health services at least once every year (Rattay *et al.*, 2013; Robert Koch-Institut, 2014). Timely access to outpatient health services makes up for the vast majority of primary and preventative care (Agency for Healthcare Research and Quality, 2018). According to census data, there are 13% of German citizens born in Germany, who are privately insured, compared to 6% of persons with a migrant background (Statistisches Bundesamt (Destatis), 2020). By weighting the estimates for appointment rates from our study with these numbers, we conclude that patients are 2.1% more likely to get an

[30]Studies across different contexts confirm that whether, and to what extent, discrimination exists, depends on the market environment. Lowande and Proctor (2020) find that public officials tasked primarily with service provision show no evidence of discrimination. This result is in contrast to Giulietti *et al.* (2019) who find strong discrimination in bureaucratic behavior and argue, however, that their finding may be due to a lack of competitive forces. Boulware and Kuttner (2019) suggest that there exists a cyclicality of discrimination that is driven by labor market tightness. Hedegaard and Tyran (2018) find in a field-experiment that workers commonly avoid working with coworkers of another ethnic type, but this behavior depends a lot on the cost of choosing a less productive worker.

appointment when they are German.[31] Noteworthy, this result is neither because of statistical nor taste-based ethnic discrimination. To equalize the access to outpatient health care services, policy-makers could enforce (web-based) system that obliges physicians to accept patients without information on their insurance type.

[31] We calculate this number as follows: estimated appointment rates of our study are 0.412 and 0.540 for SHI and PHI, respectively (see Section 4.1). The average appointment rate is then $0.94 \times 0.412 + 0.06 \times 0.540 = 42.0\%$ for patients with migrant background and $0.87 \times 0.412 + 0.13 \times 0.540 = 42.9\%$ for German patients. The increase from 41.9 to 42.9 is 2.1%.

References

Agency for Healthcare Research and Quality (2018). *National Healthcare Quality and Disparities Report*. Report, Agency for Healthcare Research and Quality, Rockville, MD, download date: April 2020.

Alsan, M., Garrick, O. and Graziani, G. (2019). Does Diversity Matter for Health? Experimental Evidence from Oakland. *American Economic Review*, **109** (12), 4071–4111.

— and Wanamaker, M. (2017). Tuskegee and the Health of Black Men. *Quarterly Journal of Economics*, **133** (1), 407–455.

Angerer, S., Waibel, C. and Stummer, H. (2019). Discrimination in Health Care: A Field Experiment on the Impact of Patients' Socioeconomic Status on Access to Care. *American Journal of Health Economics*, **5** (4), 407–427.

Antidiskriminierungsstelle des Bundes (2017). *Diskriminierung in Deutschland. Dritter Gemeinsamer Bericht der Antidiskriminierungsstelle des Bundes und der in ihrem Zuständigkeitsbereich betroffenen Beauftragten der Bundesregierung und des Deutschen Bundestages* . Tech. rep., Antidiskriminierungsstelle des Bundes, Berlin, Germany.

Arrow, K. J. (1973). The Economics of Discrimination. In O. Ashenfelter and A. Rees (eds.), *Discrimination in Labor Markets*, Princeton University Press, pp. 3–33.

Auspurg, K., Schneck, A. and Hinz, T. (2019). Closed Doors Everywhere? A Meta-Analysis of Field Experiments on Ethnic Discrimination in Rental Housing Markets. *Journal of Ethnic and Migration Studies*, **45** (1), 95–114.

Baert, S. (2017). *Hiring Discrimination: An Overview of (Almost) All Correspondence Experiments Since 2005*. IZA Discussion Paper 10738, Institute for the Study of Labor (IZA), Bonn, Germany.

Balsa, A. I. and McGuire, T. G. (2001). Statistical Discrimination in Health Care. *Journal of Health Economics*, **20**, 881–907.

Bauernschuster, S., Driva, A. and Hornung, E. (2020). Bismarck's Health Insurance and the Mortality Decline. *Journal of the European Economic Association*, **18** (5), 2561–2607.

Becker, G. S. (1957). *The Economics of Discrimination*. Chicago: University of Chicago Press.

Bertrand, M. and Duflo, E. (2017). Field Experiments on Discrimination. In A. V. Banerjee and E. Duflo (eds.), *Handbook of Economic Field Experiments*, vol. 1, *8*, Elsevier Science B.V., pp. 309–393.

Bisgaier, J. and Rhodes, K. V. (2011). Auditing Access to Specialty Care for Children with Public Insurance. *New England Journal of Medicine*, **364** (24), 2324–2333.

Blümel, M. and Busse, R. (2020). The German Health Care System. In R. Tikkanen, R. Osborn, E. Mossialos, A. Djordjevic and G. Wharton (eds.), *2020 International Profiles of Health Care Systems*, The Commonwealth Fund, pp. 83–92.

Bohren, J. A., Haggag, K., Imas, A. and Pope, D. G. (2019). *Inaccurate Statistical Discrimination*. NBER Working Paper 25935, National Bureau of Economic Research, Cambridge, MA.

Boisjoly, J., Duncan, G. J., Kremer, M., Levy, D. M. and Eccles, J. (2006). Empathy or Antipathy? The Impact of Diversity. *American Economic Review*, **96** (5), 1890–1905.

Boulware, K. D. and Kuttner, K. N. (2019). Labor Market Conditions and Discrimination: Is There a Link? *American Economic Review: Papers & Proceedings Papers and Proceedings*, **109**, 166–170.

Brekke, K. R., Siciliani, L. and Straume, O. R. (2008). Competition and Waiting Times in Hospital Markets. *Journal of Public Economics*, **92** (7), 1607–1628.

Busse, R., Blümel, M., Knieps, F. and Bärnighausen, T. (2017). Statutory Health Insurance in Germany: A Health System Shaped by 135 Years of Solidarity, Self-governance, and Competition. *The Lancet*, **390** (10097), 882–897.

Byrd, W. M. and Clayton, L. A. (2001). Race, Medicine, and Health Care in the United States: A Historical Survey. *Journal of National Medical Association*, **93** (3), 11S–34S.

Cantoni, D., Hagemeister, F. and Westcott, M. (2019). *Persistence and Activation of Right-Wing Political Ideology*. Rationality and Competition Discussion Paper Series 143, Collaborative Research Center Transregio 190.

Carol, S., Eich, D., Keller, M., Steiner, F. and Storz, K. (2019). Who Can Ride Along? Discrimination in a German Carpooling Market. *Population, Space Place*, **25** (8).

Chandra, A. and Staiger, D. O. (2010). *Identifying Provider Prejudice in Healthcare*. NBER Working Paper 16382, National Bureau of Economic Research, Cambridge, MA.

Clemens, J. and Gottlieb, J. D. (2014). Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health. *American Economic Review*, **104** (4), 1320–1349.

Cullis, J. G., Jones, P. R. and Propper, C. (2000). Waiting Lists and Medical Treatment: Analysis and Policies. In A. J. Culyer and J. P. Newhouse (eds.), *Handbook of Health Economics*, vol. 1B, Elsevier B.V., pp. 1201–1249.

Deaton, A. (2008). Income, Health, and Well-Being around the World: Evidence from the Gallup World Poll. *Journal of Economic Perspectives*, **22** (2), 53–72.

Doleac, J. L. and Stein, L. C. D. (2013). The Visible Hand: Race and Online Market Outcomes. *Economic Journal*, **123** (572), F469–F492.

Dustmann, C., ristine Vasiljeva and Damm, A. P. (2019). Refugee Migration and Electoral Outcomes. *Review of Economic Studies*, **86** (5), 2035–2091.

Edelman, B., Luca, M. and Svirsky, D. (2017). Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment. *American Economic Journal: Applied Economics*, **9** (2), 1–22.

Ewens, M., Tomlin, B. and Wang, L. C. (2014). Statistical Discrimination or Prejudice? A Large Sample Field Experiment. *Review of Economics and Statistics*, **96** (1), 119–134.

Farkas, G. (2003). Racial Disparities and Discrimination in Education: What Do We Know, How Do We Know It, and What Do We Need to Know? *Teachers College Record*, **105** (6), 1119–1146.

Fleurbaey, M. and Schokkaert, E. (2011). Equity in Health and Health Care. In M. V. Pauly, T. G. McGuire and P. P. Barros (eds.), *Handbook of Health Economics*, vol. 2, Elsevier B.V., pp. 1003–1092.

Gaynor, M. (2007). Competition and Quality in Health Care Markets. *Foundations and Trends® in Microeconomics*, **2** (6), 441–508.

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

Glover, D., Pallais, A. and Pariente, W. (2017). Discrimination as a Self-Fulfilling Prophecy: Evidence from French Grocery Stores. *Quarterly Journal of Economics*, **132** (3), 1219–1260.

Hedegaard, M. S. and Tyran, J. (2018). The Price of Prejudice. *American Economic Journal: Applied Economics*, **10** (1), 40–63.

Heinrich, N., Wübker, A. and Wuckel, C. (2018). Waiting Times for Outpatient Treatment in Germany: New Experimental Evidence from Primary Data. *Jahrbücher für Nationalökonomie und Statistik*, **238** (5), 375–394.

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

Hjort, J. (2014). Ethnic Divisions and Production in Firms. *Quarterly Journal of Economics*, **129** (4), 1899–1946.

Igel, U., Brähler, E. and Grande, G. (2010). Der Einfluss von Diskriminierungserfahrungen auf die Gesundheit von MigrantInnen. *Psychiatrische Praxis*, **37** (4), 183–190.

Institute of Medicine (IOM) (2003). *Unequal Treatment. Confronting Racial and Ethnic Disparities in Health Care.* Washington, DC: National Academies Press.

Iversen, T. (1993). A Theory of Hospital Waiting Lists. *Journal of Health Economics*, **12** (1), 55–71.

Johar, M., Jones, G., Keane, M. P., Savage, E. and Stavrunova, O. (2013). Discrimination in a Universal Health System: Explaining Socioeconomic Waiting Time Gaps. *Journal of Health Economics*, **32** (1), 181–194.

Kaas, L. and Manger, C. (2011). Ethnic Discrimination in Germany's Labour Market: A Field Experiment. *German Economic Review*, **13** (1), 1–20.

Kessler, J. B., Low, C. and Sullivan, C. D. (2019). Incentivized Resume Rating: Eliciting Employer Preferences without Deception. *American Economic Review*, **109** (11), 3713–3744.

Kuchinke, B. A., Sauerland, D. and Wübker, A. (2009). The Influence of Insurance Status on Waiting Times in German Acute Care Hospitals: An Empirical Analysis of New Data. *International Journal of Equity in Health*, **8** (44).

Kugelmass, H. (2016). "Sorry, I'm Not Accepting New Patients" An Audit Study of Access to Mental Health Care. *Journal of Health and Social Behavior*, **57** (2), 168–183.

Lowande, K. and Proctor, A. (2020). Bureaucratic Responsiveness to LGBT Americans. *American Journal of Political Science*, **64** (3), 664–681.

LUNGEN, M., STOLLENWERK, B., MESSNER, P., LAUTERBACH, K. W. and GERBER, A. (2008). Waiting Times for Elective Treatments According to Insurance Status: A Randomized Empirical Study in Germany. *International Journal for Equity in Health*, **7**, 1.

MCGUIRE, T. G. (2000). Physician Agency. In A. J. Culyer and J. P. Newhouse (eds.), *Handbook of Health Economics*, vol. 1A, Elsevier Science B.V., pp. 461–536.

MILKMAN, K. L., AKINOLA, M. and CHUGH, D. (2015). What Happens Before? A Field Experiment Exploring How Pay and Representation Differentially Shape Bias on the Pathway into Organizations. *Journal of Applied Psychology*, **100** (6), 1678–1712.

MONSTAD, K., ENGESÆTER, L. B. and ESPEHAUG, B. (2014). Waiting Time and Socioeconomic Status – An Individual-level Analysis. *Health Economics*, **23** (4), 446–461.

MOSCELLI, G., SICILIANI, L., GUTACKER, N. and COOKSON, R. (2018). Socioeconomic Inequality of Access to Healthcare: Does Choice Explain the Gradient? *Journal of Health Economics*, **57**, 290–314.

MUJCIC, R. and FRIJTERS, P. (2021). The Colour of a Free Ride. *Economic Journal*, **131** (634), 970–999.

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

PASCOE, E. A. and RICHMAN, L. S. (2009). Perceived Discrimination and Health: A Meta-Analytic Review. *Psychological Bulletin*, **135** (4), 531–554.

PHELPS, E. S. (1972). The Statistical Theory of Racism and Sexism. *American Economic Review*, **62** (4), 659–661.

POLSKY, D., RICHARDS, M., BASSEYN, S., WISSOKER, D., KENNEY, G. M., ZUCKERMAN, S. and RHODES, K. V. (2015). Appointment Availability after Increases in Medicaid Payments for Primary Care. *New England Journal of Medicine*, **372** (6), 537–545.

RATTAY, P., BUTSCHALOWSKY, H., ROMMEL, A., PRÜTZ, F., JORDAN, S., NOWOSSADECK, E., DOMANSKA, O. and KAMTSIURIS, P. (2013). Inanspruchnahme der ambulanten und stationären medizinischen Versorgung in Deutschland. *Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz*, **56**, 832–844.

RIACH, P. A. and RICH, J. (2002). Field Experiments of Discrimination in the Market Place. *Economic Journal*, **112** (483), F480–F518.

Robert Koch-Institut (2014). *Daten und Fakten: Ergebnisse der Studie "Gesundheit in Deutschland aktuell 2012"*. Beiträge zur gesundheitsberichterstattung des bundes, Robert Koch-Institut, Abteilung für Epidemiologie und Gesundheitsmonitoring, Berlin, Germany.

Schührer, S. (2018). *Türkeistämmige Personen in Deutschland. Erkenntnisse aus der Repräsentativuntersuchung Ausgewählte Migrantengruppen in Deutschland 2015*. Working Paper des Forschungszentrums des Bundesamtes 81, Bundesamt für Migration und Flüchtlinge, Nürnberg, Germany.

Schwierz, C., Wübker, A., Wübker, A. and Kuchinke, B. A. (2011). Discrimination in Waiting Times by Insurance Type and Financial Soundness of German Acute Care Hospitals. *European Journal of Health Economics*, **12** (5), 405–416.

Sharma, R., Mitra, A. and Stano, M. (2015). Insurance, Race/Ethnicity, and Sex in the Search for a New Physician. *Economics Letters*, **137**, 150–153.

Siciliani, L., Moran, V. and Borowitz, M. (2014). Measuring and Comparing Health Care Waiting Times in OECD Countries. *Health Policy*, **118** (3), 1201–1249.

Statistisches Bundesamt (Destatis) (2019). *Bevölkerung und Erwerbstätigkeit. Ausländische Bevölkerung. Ergebnisse des Ausländerzentralregisters 2018*. Fachserie 1, reihe 2, Statistisches Bundesamt (Destatis), Wiesbaden, Germany.

Statistisches Bundesamt (Destatis) (2020). *Sozialleistungen - Angaben zur Krankenversicherung (Ergebnisse des Mikrozensus 2019)*. Tech. rep., Statistisches Bundesamt (Destatis), Wiesbaden, Germany.

Stefan, M., Holzmeister, F., Müllauer, A. and Kirchler, M. (2018). Ethnical Discrimination in Europe: Field Evidence from the Finance Industry. *PLoS One*, **13** (1).

Sundmacher, L. and Ozegowski, S. (2016). Regional Distribution of Physicians: The Role of Comprehensive Private Health Insurance in Germany. *European Journal of Health Economics*, **17** (4), 443–451.

Verband der Privaten Krankenversicherung (2020). *Zahlenbericht 2019*. Tech. rep., Verband der privaten Krankenversicherung e.V., Köln, Germany.

Wagstaff, A. and Van Doorslaer, E. (2000). Measuring and Testing for Inequity in the Delivery of Health Care. *Journal of Human Resources*, **35** (4), 716–733.

Walendzik, A., Manouguian, M., Gress, S. and Wasem, J. (2009). Vergütungsunterschiede im ambulanten ärztlichen Bereich zwischen PKV und GKV und Modelle der Vergütungsangleichung. *Sozialer Fortschritt*, **58** (4), 63–69.

WEICHSELBAUMER, D. (2020). Multiple Discrimination against Female Immigrants Wearing Headscarves. *ILR Review*, **73** (3), 600–627.

WERBECK, A., WÜBKER, A. and ZIEBARTH, N. R. (2021). Cream Skimming by Health Care Providers and Inequality in Health Care Access: Evidence from a Randomized Field Experiment. *Journal of Economic Behavior & Organization*, **188**, 1325–1350.

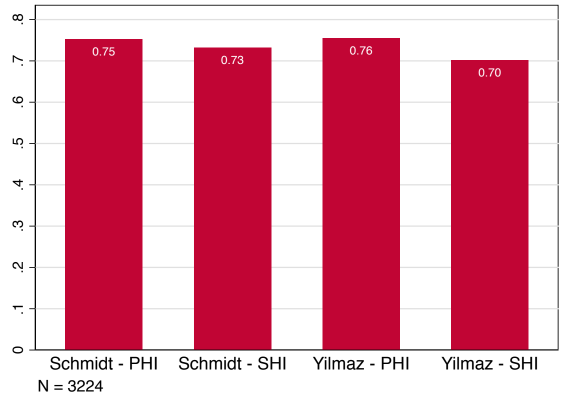
WISNIEWSKI, J. M. and WALKER, B. (2020). Association of Simulated Patient Race/Ethnicity with Scheduling of Primary Care Appointments. *JAMA Network Open*, **3** (1), e1920010–e1920010.

ZUSSMAN, A. (2013). Ethnic Discrimination: Lessons from the Israeli Online Market for Used Cars. *The Economic Journal*, **123** (572), F433–F468.

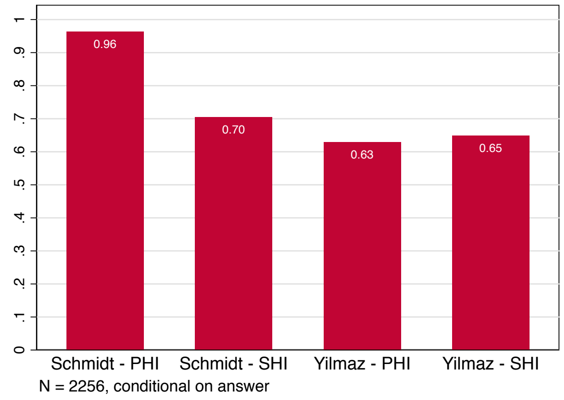
6 Figures (to be placed in the paper)

Figure 1: Averages outcomes by type of fictitious patient

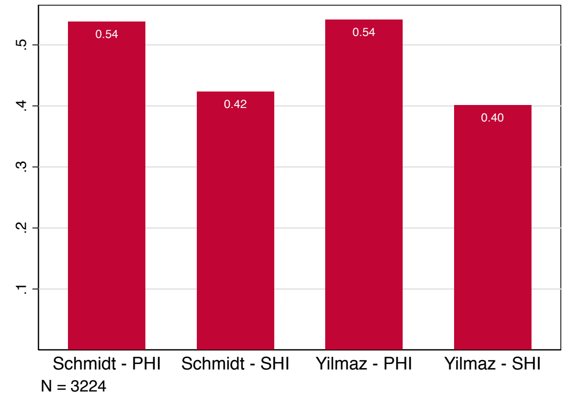
(a) Share who receive an email



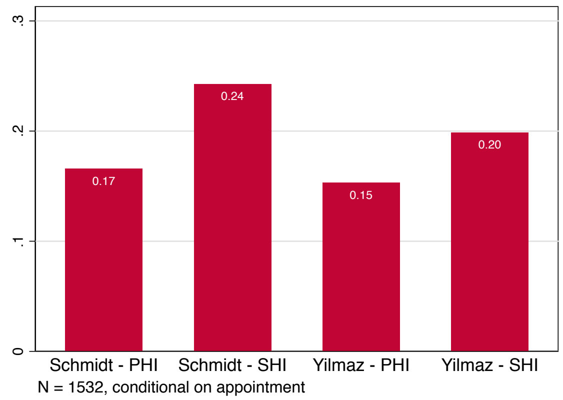
(b) Average number of days to response



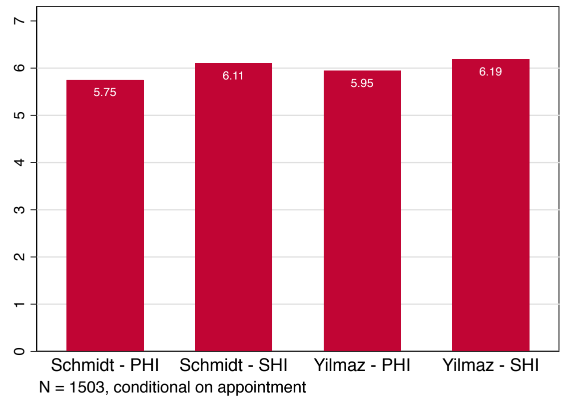
(c) Share who receive an appointment



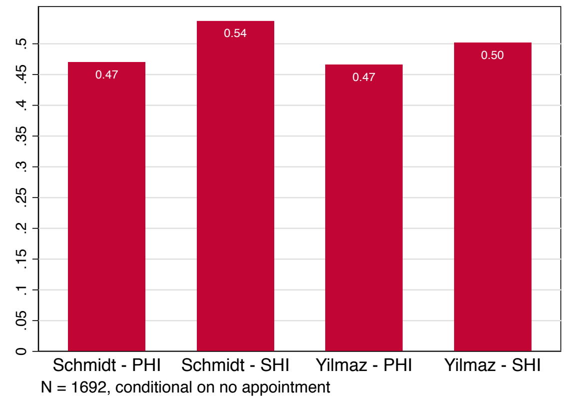
(d) Share who receive an appointment with wait



(e) Average number of days to appointment

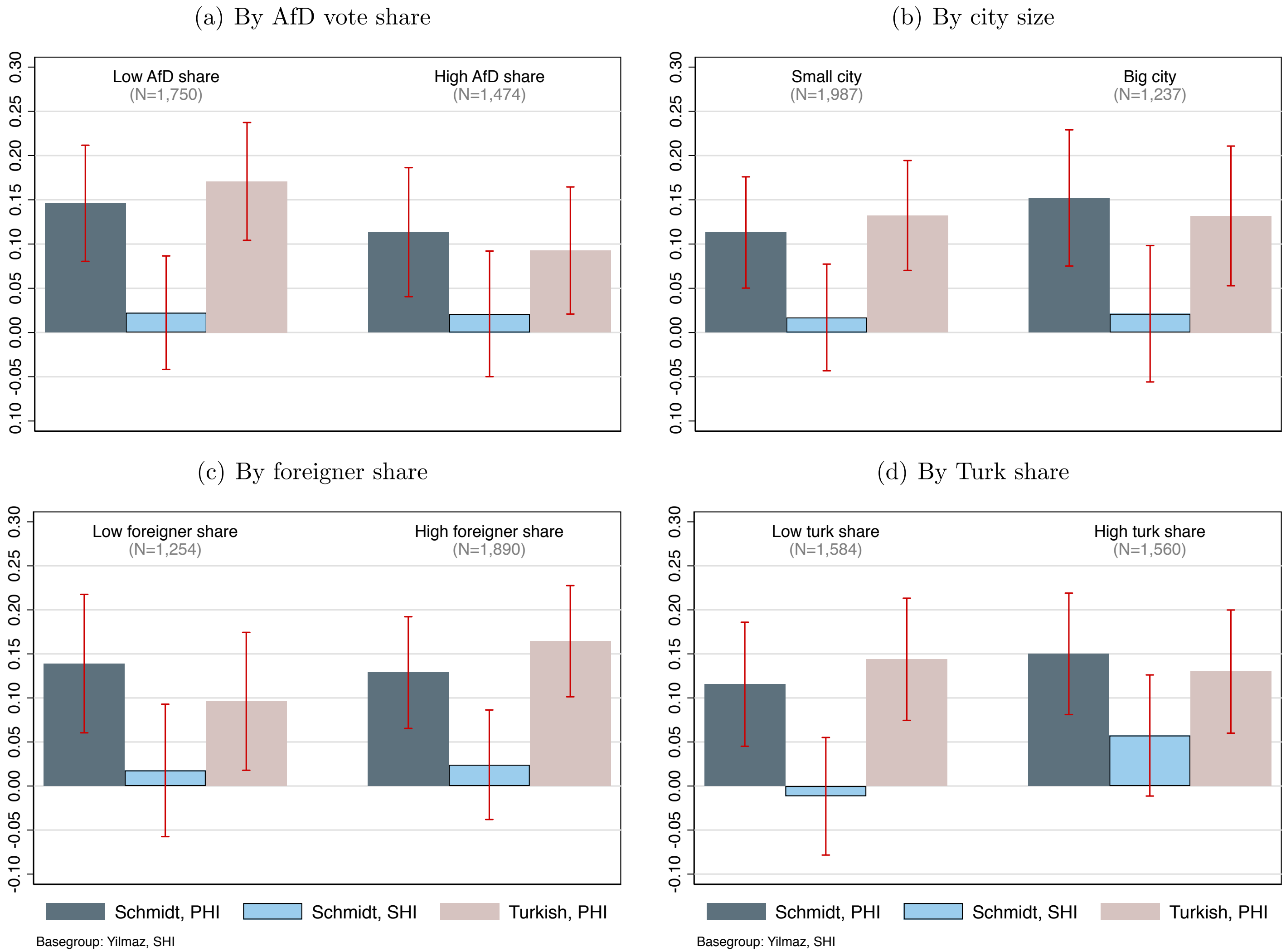


(f) Share without appointment, but response



Notes: Each panel shows the average of an outcome variable by type of fictitious patient (randomized treatment). PHI stands for private health insurance. SHI stands for statutory health insurance. Schmidt indicates a German patient, while Yilmaz a Turkish. Appendix Figure A.3 shows kernel density estimates for our two non-binary outcomes from panel (b) and (e).

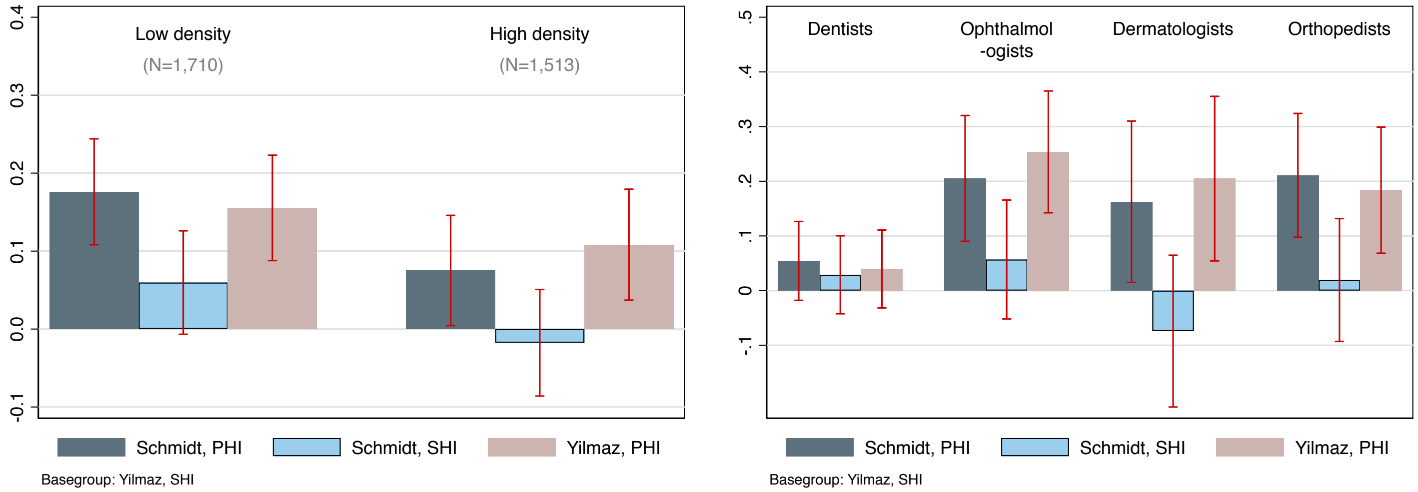
Figure 2: Estimated effect of discrimination in the likelihood of an appointment in different samples (part I)



Notes: These figures summarize estimation results from OLS regressions for the binary outcome variable appointment. In each panel, we distinguish between two different sub-samples. PHI stands for private health insurance. SHI stands for statutory health insurance. Schmidt indicates a German patient, while Yilmaz a Turkish. The base group is equal to Yilmaz, SHI. 95 percent confidence intervals (indicated in red) are based on robust standard errors. Covariates comprise physician's type fixed-effects and city fixed-effects. Appendix Tables A.2 to A.5 provides more detailed estimation output for this, and all other outcomes.

Figure 3: Estimated effect of discrimination in the likelihood of an appointment in different samples (part II)

(a) By physician density (b) By type of physician



Notes: These figures summarize estimation results from OLS regressions for the binary outcome variable appointment. In each panel, we distinguish between two different sub-samples. PHI stands for private health insurance. SHI stands for statutory health insurance. Schmidt indicates a German patient, while Yilmaz a Turkish. The base group is equal to Yilmaz, SHI. 95 percent confidence intervals (indicated in red) are based on robust standard errors. Covariates comprise physician's type fixed-effects and city fixed-effects. Appendix Tables A.6 to A.7 provides more detailed estimation output for this, and all other outcomes.

7 Tables (to be placed in the paper)

Table 1: Covariates by type of fictitious patient

	Schmidt SHI ($N = 796$)	Yilmaz SHI ($N = 817$)	Schmidt PHI ($N = 815$)	Yilmaz PHI ($N = 796$)
Physician's type				
Dentist	0.47	0.48	0.49	0.48
Ophthalmologist	0.19	0.21	0.18	0.21
Dermatologist	0.12	0.12	0.13	0.14
Orthopedist	0.21	0.19	0.20	0.17
Other physician's characteristics				
Group practice	0.41	0.44	0.42	0.41
Weekly office hours	34.66	34.49	34.87	34.49
	(10.59)	(10.67)	(11.25)	(10.82)
Office hours imputed	0.11	0.09	0.09	0.12
Physician's rank on Yellow pages	64.29	55.63	67.30	59.00
	(126.99)	(119.41)	(150.31)	(125.29)
Email sent in				
Wave 1	0.21	0.23	0.21	0.23
Wave 2	0.79	0.77	0.78	0.77

Notes: Columns shows the average (and in parentheses standard deviations) of covariates by fictitious patients. SHI stands for statutory health insurance. PHI stands for private health insurance. Schmidt indicates a German patient, while Yilmaz a Turkish.

Table 2: Average outcomes by type of physician

	Overall ($N = 3{,}224$)	Dentist ($N = 1{,}552$)	Ophthal -mologist ($N = 638$)	Dermatol -ogist ($N = 405$)	Ortho -pedist ($N = 629$)
Response rate	0.74	0.74	0.75	0.74	0.70
Days to response†	0.74	0.66	0.65	0.92	0.91
Appointment rate	0.48	0.53	0.46	0.42	0.40
Long wait‡	0.19	0.07	0.52	0.28	0.12
Time to appointment‡	5.96	6.09	5.65	5.88	5.94
Declined, but answered¶	0.50	0.46	0.53	0.56	0.51

Notes: This table lists the average of our six outcome variables in the overall sample, and by type of physician. † Conditional on response (N=2,256). ‡ Conditional on appointment (N=1,532). ¶ Conditional on no appointment (N=1,532). Appendix Figure A.1 provides an equivalent graphical depiction of these figures.

Table 3: Baseline estimation results

	Answer (yes/no)		Time to answer (in days)		Appointment (yes/no)		Appointment with long wait		Time to app. (in days)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	No covs	Covs	No covs	Covs	No covs	Covs	No covs	Covs	No covs	Covs
Schmidt, PHI	0.051**	0.042*	0.315***	0.314***	0.136***	0.131***	−0.033	−0.063**	−0.439*	−0.402*
	(0.022)	(0.022)	(0.108)	(0.109)	(0.025)	(0.025)	(0.029)	(0.026)	(0.232)	(0.236)
Schmidt, SHI	0.030	0.024	0.056	0.047	0.022	0.018	0.044	0.017	−0.083	−0.095
	(0.022)	(0.023)	(0.097)	(0.094)	(0.024)	(0.024)	(0.032)	(0.028)	(0.250)	(0.250)
Yilmaz, PHI	0.053**	0.048**	−0.019	−0.025	0.140***	0.133***	−0.046	−0.073***	−0.243	−0.211
	(0.022)	(0.022)	(0.089)	(0.091)	(0.025)	(0.025)	(0.028)	(0.026)	(0.237)	(0.241)
Constant	0.702***	0.696***	0.649***	0.509***	0.401***	0.576***	0.199***	0.263***	6.189***	5.034***
	(0.016)	(0.050)	(0.071)	(0.191)	(0.017)	(0.057)	(0.022)	(0.061)	(0.190)	(0.511)
Controls‡	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
No. of observations	3,224	3,224	2,256	2,256	3,224	3,224	1,532	1,532	1,503	1,503
Mean of outcome§	0.702	0.702	0.649	0.649	0.401	0.401	0.199	0.199	6.189	6.189
R-squared	0.002	0.040	0.007	0.058	0.016	0.065	0.008	0.283	0.003	0.097

Notes: This tables summarizes estimation results from OLS regressions for five outcome variables. Per outcome variable two specifications are shown. SHI stands for statutory insurance. PHI stands for private insurance. Robust standard errors are reported in parenthesis. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. † The base group (=constant) is equal to Yilmaz, SHI. ‡ Covariates comprise doctor's offices's rank on Gelbe Seiten, a binary indicator for group practice, the total weekly office hours, a binary indicator for imputed office hours, physician's type fixed-effects, and city fixed-effects. § Refers to base group.

Table 4: Density of physicians, by type

Type	Physicians per 100.000 inhabitants			
	Mean	Standard deviation	Minimum	Maximum
Dentists	66.97	12.89	40.77	91.48
Dermatologists	5.85	1.70	2.70	10.50
Ophthalmologists	8.87	2.2	5.09	17.95
Orthopedists	8.72	2.3	4.82	14.60

Notes: These descriptive statistics are based on city-level data ($N = 79$). Densities are defined as the number physicians per 100,000 inhabitants. The data on dentists refer to the year 2017 and are provided by the *Kassenzahnärztliche Bundesvereinigung* (sent via email on 20. November 2019). The other data refer to the year 2013 and are provided by the *Zentralinstitut für die kassenärztliche Versorgung in der Bundesrepublik Deutschland* (downloaded via https://www.versorgungsatlas.de/). Appendix Figure A.4 shows Kernel density estimates.

Web Appendix

This Web Appendix (not for publication) provides additional material discussed in the unpublished manuscript "Testing for Ethnic Discrimination in Outpatient Health Care: Evidence from a Field Experiment in Germany" by Martin Halla, Christopher Kah, and Rupert Sausgruber.

A.1 List of cities in sample

We contacted physicians from the 79 largest cities in Germany. This list of cities is form largest to smallest as follows: Berlin, Hamburg, München, Köln, Frankfurt am Main, Stuttgart, Düsseldorf, Dortmund, Essen, Leipzig, Bremen, Dresden, Hannover, Nürnberg, Duisburg, Bochum, Wuppertal, Bielefeld, Bonn, Münster, Karlsruhe, Mannheim, Augsburg, Wiesbaden, Gelsenkirchen, Mönchengladbach, Braunschweig, Chemnitz, Kiel, Aachen, Halle (Saale), Magdeburg, Freiburg im Breisgau, Krefeld, Lübeck, Oberhausen, Erfurt, Mainz, Rostock, Kassel, Hagen, Hamm, Saarbrücken, Mülheim an der Ruhr, Potsdam, Ludwigshafen am Rhein, Oldenburg, Leverkusen, Osnabrück, Solingen, Heidelberg, Herne, Neuss, Darmstadt, Paderborn, Regensburg, Ingolstadt, Würzburg, Fürth, Wolfsburg, Offenbach am Main, Ulm, Heilbronn, Pforzheim, Göttingen, Bottrop, Trier, Recklinghausen, Reutlingen, Bremerhaven, Koblenz, Bergisch Gladbach, Jena, Remscheid, Erlangen, Moers, Siegen, Hildesheim, and Salzgitter.

A.2 Wording of correspondence

A.2.1 Original German wording of correspondence

Dentist

Sehr geehrte Damen & Herren,

ich bräuchte dringend einen Termin, da eine Füllung locker ist. Ich bin gerade im Urlaub und komme am Montag zurück nach Hause. Ich bin bei der [Name der Versicherung] versichert. Wann wäre der nächstmögliche Termin?

Vielen Dank im Voraus!

Mit freundlichen Grüßen

[NAME]

Ophthalmologist

Sehr geehrte Damen & Herren,

ich bräuchte dringend einen Termin, da ich ein entzündetes Auge habe. Ich bin gerade im Urlaub und komme am Montag zurück nach Hause. Ich bin bei der [Name der Versicherung] versichert. Wann wäre der nächstmögliche Termin?

Vielen Dank im Voraus!

Mit freundlichen Grüßen

[NAME]

Orthopedists

Sehr geehrte Damen & Herren,

ich bräuchte dringend einen Termin, da ich Rückenbeschwerden habe. Ich bin gerade im Urlaub und komme am Montag zurück nach Hause. Ich bin bei der [Name der Versicherung] versichert. Wann wäre der nächstmögliche Termin?

Vielen Dank im Voraus!

Mit freundlichen Grüßen

[NAME]

Dermatologists

Sehr geehrte Damen & Herren,

ich bräuchte dringend einen Termin, da ich einen Ausschlag im Genitalbereich habe. Ich bin gerade im Urlaub und komme am Montag zurück nach Hause. Ich bin bei der [Name der Versicherung] versichert. Wann wäre der nächstmögliche Termin?

Vielen Dank im Voraus!

Mit freundlichen Grüßen

[NAME]

A.2.2 Translated wording of correspondence

Dentist

Dear Sir or Madam,

I urgently need an appointment because a filling is loose. I'm on vacation right now and I'll be back home on Monday. I am insured with [name of insurance company]. When would be the next possible date?

Thanks in advance!

Kind regards

[NAME]

Ophthalmologist

Dear Sir or Madam,

I urgently need an appointment because I have an inflammation in my eye. I'm on vacation right now and I'll be back home on Monday. I am insured with [name of insurance company]. When would be the next possible date?

Thanks in advance!

Kind regards

[NAME]

Orthopedists

Dear Sir or Madam,

I urgently need an appointment because I have back pain. I'm on vacation right now and I'll be back home on Monday. I am insured with [name of insurance company]. When would be the next possible date?

Thanks in advance!

Kind regards

[NAME]

Dermatologists

Dear Sir or Madam,

I urgently need an appointment because I have a rash in my genital area. I'm on vacation right now and I'll be back home on Monday. I am insured with [name of insurance company]. When would be the next possible date?

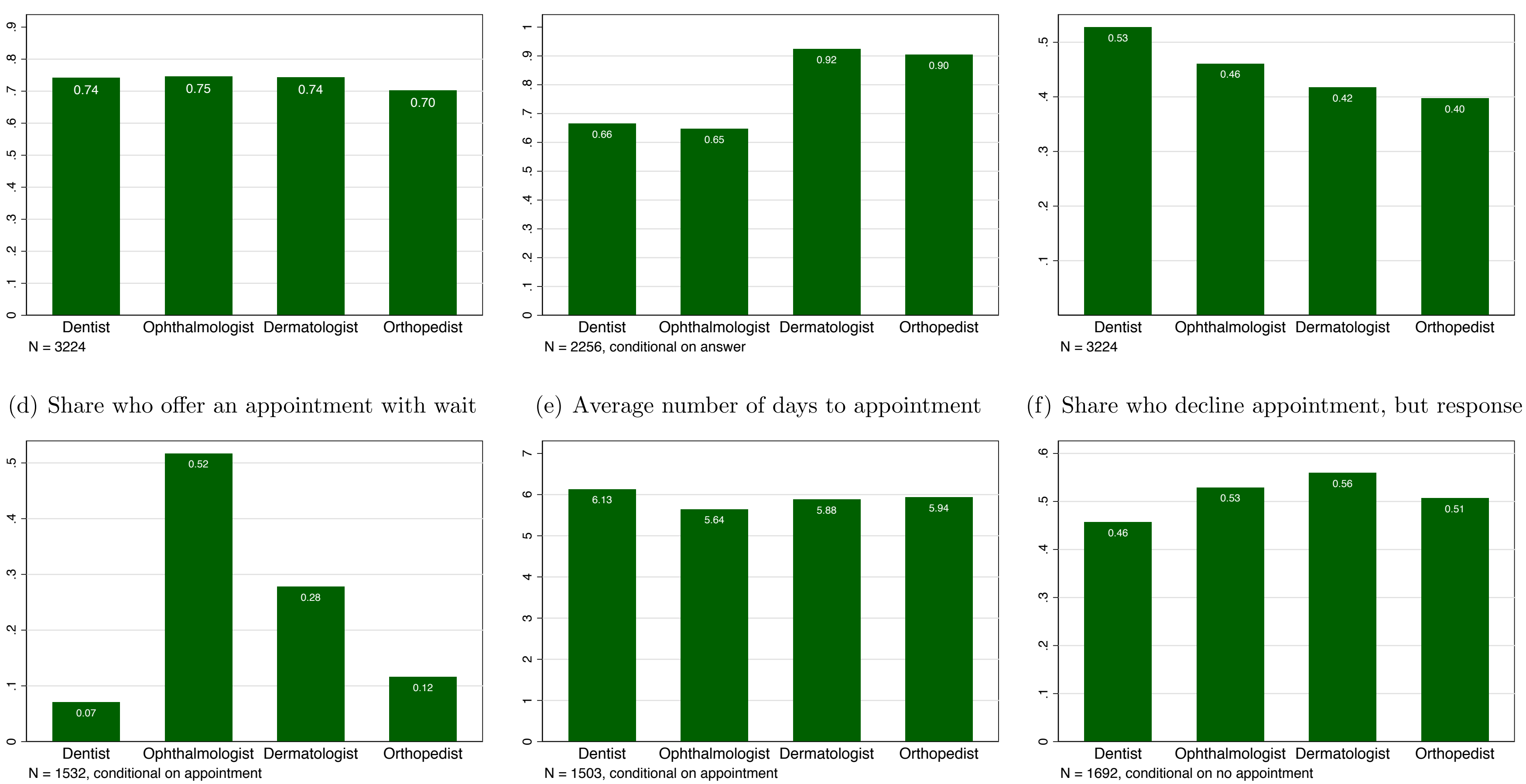
Thanks in advance!

Kind regards

[NAME]

A.3 Additional Figures

Figure A.1: Averages outcomes by type of physician



(a) Share who answer email

(b) Average number of days to response

(c) Share who offer an appointment

(d) Share who offer an appointment with wait

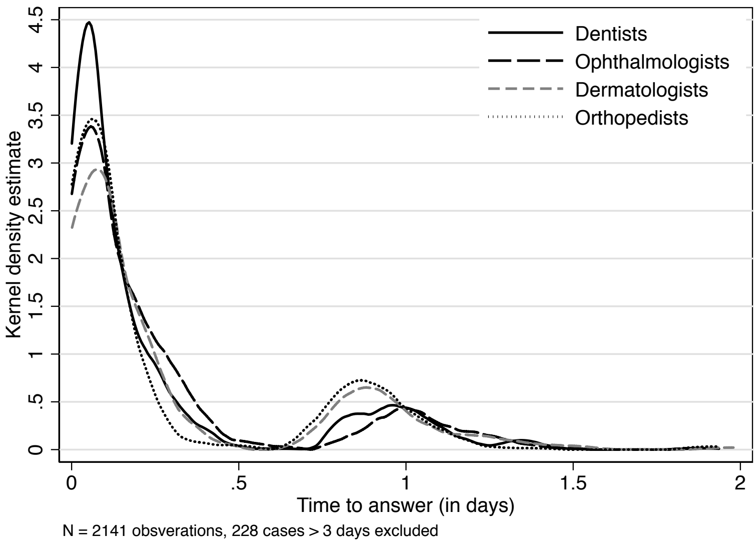
(e) Average number of days to appointment

(f) Share who decline appointment, but response

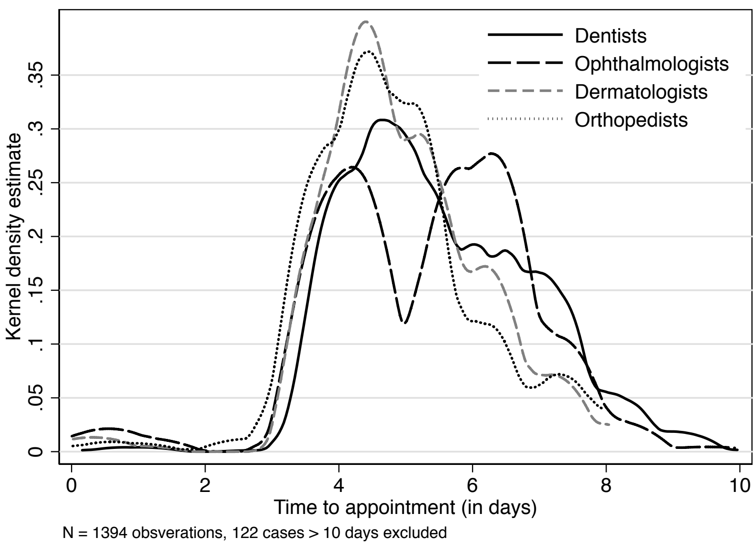
Notes: Each panel shows the average of an outcome variable by type of physician.

Figure A.2: Kernel density estimates of non-binary outcomes by type of physician

(a) Number of days to response



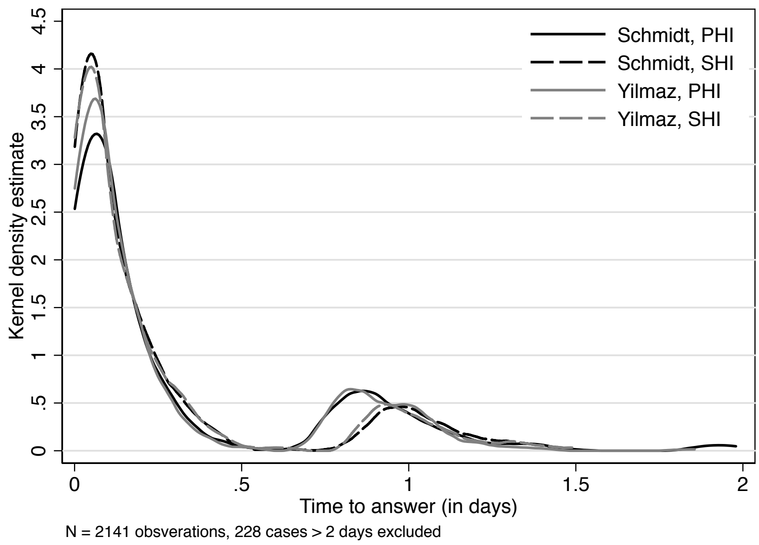
(b) Number of days to appointment



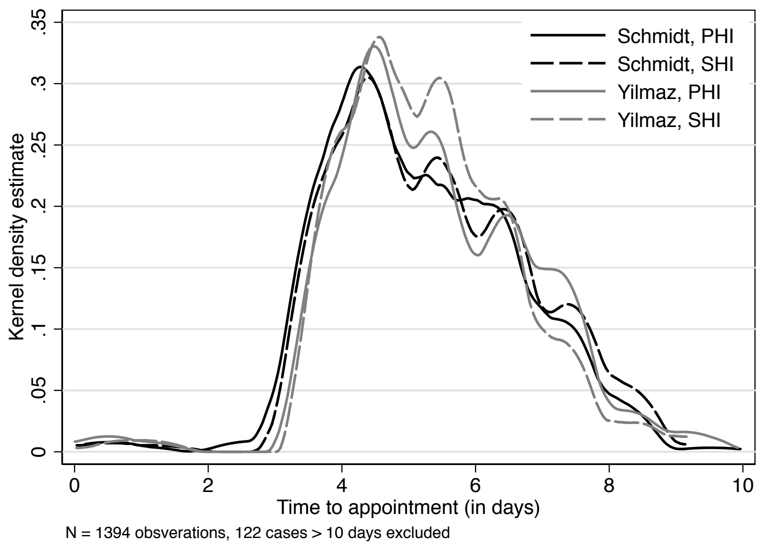
Notes: These figures show kernel density estimates for our two non-binary outcomes.

Figure A.3: Kernel density estimates of non-binary outcomes by type of fictitious patient

(a) Number of days to response



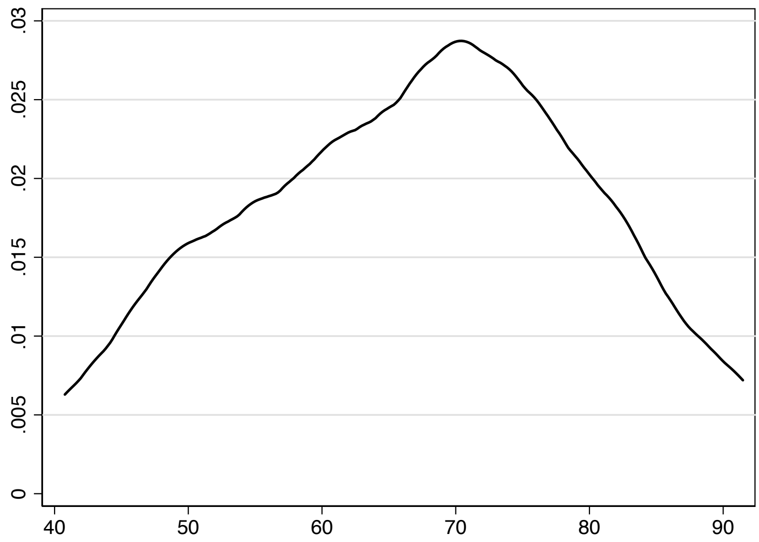
(b) Number of days to appointment



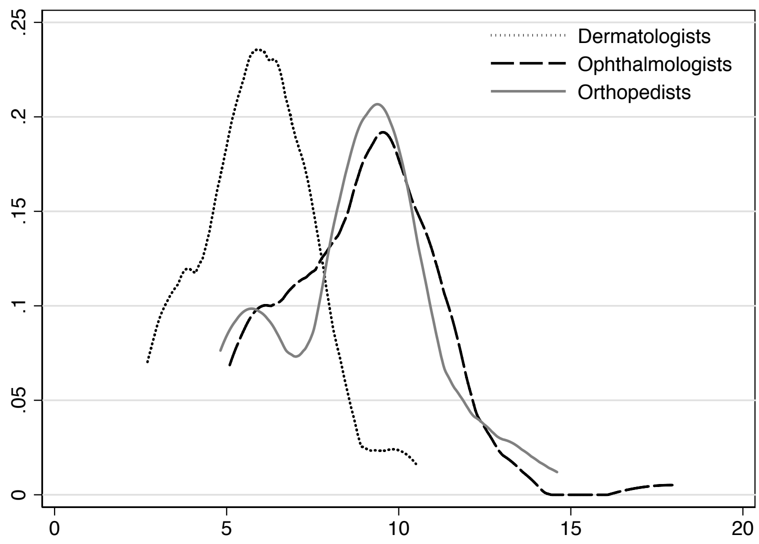
Notes: These figures show kernel density estimates for our two non-binary outcomes. PHI stands for private health insurance. SHI stands for statutory health insurance. Schmidt indicates a German patient, while Yilmaz a Turkish.

Figure A.4: Kernel estimates of physicians densities

(a) Dentists



(b) Other types of physicians



Notes: These Kernel estimates of physicians densities across German cities are based on city-level data ($N = 79$). Densities are defined as the number physicians per 100,000 inhabitants. The data on dentists refer to the year 2017 and are provided by the *Kassenzahnärztliche Bundesvereinigung* (sent via email on 20. November 2019). The other data refer to the year 2013 and are provided by the *Zentralinstitut für die kassenärztliche Versorgung in der Bundesrepublik Deutschland* (downloaded via `https://www.versorgungsatlas.de/`, add date).

A.4 Additional Tables

Table A.1: Estimated effects on the likelihood of an appointment, various specifications

	(1)	(2)	(3)	(4)
Schmidt, PHI	0.136***	0.139***	0.131***	0.113***
	(0.025)	(0.025)	(0.025)	(0.025)
Schmidt, SHI	0.022	0.023	0.018	
	(0.024)	(0.024)	(0.024)	
Yilmaz, PHI	0.140***	0.141***	0.133***	0.115***
	(0.025)	(0.025)	(0.025)	(0.025)
Yilmaz, SHI				−0.018
				(0.024)
Constant†	0.401***	0.453***	0.576***	0.594***
	(0.017)	(0.020)	(0.057)	(0.057)
Doctor's rank on Yellow pages			−0.000**	−0.000**
			(0.000)	(0.000)
Group practice			0.034*	0.034*
			(0.019)	(0.019)
Total office hours per week			0.000	0.000
			(0.001)	(0.001)
Office hours imputed			−0.091***	−0.091***
			(0.029)	(0.029)
Physician's type FE	No	Yes	Yes	Yes
City FE	No	No	Yes	Yes
No. of observations	3,224	3,224	3,224	3,224
Mean of outcome‡	0.401	0.401	0.401	0.424
R-squared	0.016	0.029	0.065	0.065

Notes: This tables summarizes estimation results from OLS regressions for the binary outcome variable indicating the offer of an appointment. There are three specifications with varying covariates. SHI stands for statutory insurance. PHI stands for private insurance. Robust standard errors are reported in parenthesis. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. † In columns (1) to (3) the base group (=constant) is equal to Yilmaz, SHI. In column (4) he base group (=constant) is equal to Schmidt, SHI. ‡ Refers to base group.

Table A.2: Estimation results by AfD vote share in 2017

	Answer (yes/no)		Time to answer (in days)		Appointment (yes/no)		Appointment with long wait		Time to app. (in days)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Low	High	Low	High	Low	High	Low	High	Low	High
Schmidt, PHI	0.064**	0.017	0.366**	0.245	0.146***	0.113***	−0.090***	−0.035	−0.338	−0.406
	(0.030)	(0.033)	(0.156)	(0.151)	(0.033)	(0.037)	(0.033)	(0.041)	(0.313)	(0.355)
Schmidt, SHI	0.058*	−0.011	0.062	0.024	0.022	0.021	0.034	0.007	−0.160	−0.065
	(0.030)	(0.034)	(0.130)	(0.139)	(0.033)	(0.036)	(0.037)	(0.044)	(0.329)	(0.379)
Yilmaz, PHI	0.080***	0.016	−0.131	0.102	0.171***	0.093**	−0.092***	−0.048	−0.450	0.109
	(0.030)	(0.033)	(0.121)	(0.137)	(0.034)	(0.037)	(0.033)	(0.041)	(0.304)	(0.390)
Constant†	0.646***	0.743***	0.563*	0.386***	0.537***	0.536***	0.217	0.146***	6.268***	5.711***
	(0.099)	(0.038)	(0.335)	(0.135)	(0.098)	(0.042)	(0.132)	(0.045)	(0.972)	(0.421)
Covariates‡	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	1,750	1,474	1,245	1,011	1,750	1,474	843	689	827	676
Mean of outcome§	0.578	0.628	0.622	0.679	0.397	0.406	0.184	0.216	6.271	6.098
R-squared	0.037	0.033	0.057	0.054	0.067	0.053	0.264	0.298	0.105	0.101

Notes: This tables summarizes estimation results from OLS regressions for five outcome variables, in two different subsamples (cities with a low versus high vote share for Afd in general elections in 2017). SHI stands for statutory insurance. PHI stands for private insurance. Robust standard errors are reported in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. † The base group (=constant) is equal to Yilmaz, SHI. ‡ Covariates comprise physician's type fixed-effects, and city fixed-effects. § Refers to base group.

Table A.3: Estimation results by city size

	Answer (yes/no)		Time to answer (in days)		Appointment (yes/no)		Appointment with long wait		Time to app. (in days)	
	(1) Small	(2) Big	(3) Small	(4) Big	(5) Small	(6) Big	(7) Small	(8) Big	(9) Small	(10) Big
Schmidt, PHI	0.060**	0.012	0.202	0.464***	0.113***	0.152***	−0.075**	−0.048	−0.460	−0.275
	(0.030)	(0.035)	(0.140)	(0.172)	(0.032)	(0.039)	(0.031)	(0.045)	(0.311)	(0.362)
Schmidt, SHI	0.039	0.002	0.099	−0.061	0.017	0.021	−0.006	0.064	−0.222	0.053
	(0.029)	(0.036)	(0.131)	(0.125)	(0.031)	(0.039)	(0.034)	(0.050)	(0.327)	(0.384)
Yilmaz, PHI	0.062**	0.028	−0.135	0.124	0.132***	0.132***	−0.057*	−0.093**	−0.413	0.097
	(0.029)	(0.035)	(0.123)	(0.133)	(0.032)	(0.040)	(0.031)	(0.044)	(0.301)	(0.395)
Constant†	0.696***	0.759***	1.322***	0.328**	0.369***	0.561***	0.035	0.189***	7.734***	6.212***
	(0.105)	(0.038)	(0.490)	(0.140)	(0.120)	(0.045)	(0.028)	(0.047)	(1.161)	(0.414)
Covariates‡	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	1,987	1,237	1,360	896	1,987	1,237	915	617	897	606
Mean of outcome§	0.666	0.504	0.714	0.549	0.396	0.409	0.172	0.242	6.503	5.668
R-squared	0.040	0.020	0.074	0.034	0.061	0.059	0.327	0.228	0.113	0.049

Notes: This tables summarizes estimation results from OLS regressions for five outcome variables, in two different subsamples (small versus big cities). SHI stands for statutory insurance. PHI stands for private insurance. Robust standard errors are reported in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. † The base group (=constant) is equal to Yilmaz, SHI. ‡ Covariates comprise physician's type fixed-effects, and city fixed-effects. § Refers to base group.

Table A.4: Estimation results by share of foreigners

	Answer (yes/no)		Time to answer (in days)		Appointment (yes/no)		Appointment with long wait		Time to app. (in days)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Low share	High share	Low share	High share	Low share	High share	Low share	High share	Low share	High share
Schmidt, PHI	0.063*	0.026	0.213	0.394***	0.139***	0.129***	−0.061	−0.055	−0.430	−0.343
	(0.036)	(0.029)	(0.189)	(0.130)	(0.040)	(0.032)	(0.041)	(0.035)	(0.374)	(0.317)
Schmidt, SHI	0.015	0.035	−0.057	0.102	0.018	0.024	0.024	0.028	−0.110	−0.129
	(0.036)	(0.029)	(0.160)	(0.114)	(0.038)	(0.032)	(0.043)	(0.038)	(0.413)	(0.327)
Yilmaz, PHI	0.013	0.080***	−0.194	0.094	0.096**	0.164***	−0.086**	−0.059*	−0.658*	0.074
	(0.038)	(0.028)	(0.158)	(0.109)	(0.040)	(0.032)	(0.042)	(0.033)	(0.367)	(0.326)
Constant†	0.700***	0.726***	0.187	0.271**	0.434***	0.509***	0.307*	0.199***	6.161***	5.996***
	(0.102)	(0.035)	(0.162)	(0.124)	(0.113)	(0.039)	(0.179)	(0.042)	(1.031)	(0.384)
Covariates‡	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	1,254	1,890	874	1,323	1,254	1,890	600	894	588	878
Mean of outcome§	0.676	0.544	0.721	0.588	0.418	0.390	0.192	0.196	6.489	6.002
R-squared	0.039	0.036	0.070	0.047	0.092	0.047	0.319	0.256	0.110	0.086

Notes: This tables summarizes estimation results from OLS regressions for five outcome variables, in two different subsamples (low versus high share of foreigners). Information on the share of foreigners is missing for three cities (Bergisch Gladbach, Moers, Saarbruecken). SHI stands for statutory insurance. PHI stands for private insurance. Robust standard errors are reported in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. † The base group (=constant) is equal to Yilmaz, SHI. ‡ Covariates comprise physician's type fixed-effects, and city fixed-effects. § Refers to base group.

Table A.5: Estimation results by share of Turks

	Answer (yes/no)		Time to answer (in days)		Appointment (yes/no)		Appointment with long wait		Time to app. (in days)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Low share	High share	Low share	High share	Low share	High share	Low share	High share	Low share	High share
Schmidt, PHI	0.029	0.052	0.233	0.409***	0.116***	0.150***	−0.033	−0.072*	−0.482	−0.267
	(0.033)	(0.032)	(0.163)	(0.146)	(0.036)	(0.035)	(0.038)	(0.037)	(0.342)	(0.347)
Schmidt, SHI	0.008	0.047	0.066	0.019	−0.012	0.057	0.039	0.020	−0.137	−0.092
	(0.032)	(0.033)	(0.141)	(0.120)	(0.034)	(0.035)	(0.040)	(0.041)	(0.372)	(0.358)
Yilmaz, PHI	0.042	0.064**	−0.149	0.128	0.144***	0.130***	−0.086**	−0.046	−0.583*	0.199
	(0.032)	(0.032)	(0.131)	(0.124)	(0.035)	(0.036)	(0.035)	(0.038)	(0.327)	(0.372)
Constant†	0.814***	0.720***	2.294***	0.340***	0.324***	0.493***	0.106	0.172***	7.707***	5.994***
	(0.086)	(0.037)	(0.733)	(0.123)	(0.098)	(0.041)	(0.081)	(0.043)	(1.168)	(0.388)
Covariates‡	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	1,584	1,560	1,111	1,086	1,584	1,560	748	746	735	731
Mean of outcome§	0.656	0.534	0.702	0.577	0.411	0.390	0.170	0.219	6.527	5.859
R-squared	0.034	0.039	0.065	0.047	0.081	0.050	0.240	0.315	0.112	0.076

Notes: This tables summarizes estimation results from OLS regressions for five outcome variables, in two different subsamples (low versus high share of Turks). The share of Turks is missing for 3 cities (Bergisch Gladbach, Moers, Saarbruecken). SHI stands for statutory insurance. PHI stands for private insurance. Robust standard errors are reported in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. † The base group (=constant) is equal to Yilmaz, SHI. ‡ Covariates comprise physician's type fixed-effects, and city fixed-effects. § Refers to base group.

Table A.6: Estimation results by physician's type-specific density

	Answer (yes/no)		Time to answer (in days)		Appointment (yes/no)		Appointment with long wait		Time to app. (in days)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Low	High	Low	High	Low	High	Low	High	Low	High
Schmidt, PHI	0.069**	0.010	0.459***	0.188	0.176***	0.075**	−0.086**	−0.037	−0.224	−0.543
	(0.031)	(0.033)	(0.153)	(0.153)	(0.035)	(0.036)	(0.040)	(0.033)	(0.325)	(0.363)
Schmidt, SHI	0.036	0.013	0.178	−0.066	0.060*	−0.018	0.037	0.003	−0.039	−0.257
	(0.031)	(0.033)	(0.126)	(0.134)	(0.034)	(0.035)	(0.044)	(0.035)	(0.357)	(0.378)
Yilmaz, PHI	0.075**	0.020	0.183	−0.202	0.155***	0.108***	−0.086**	−0.057*	−0.132	−0.298
	(0.030)	(0.034)	(0.125)	(0.132)	(0.034)	(0.036)	(0.040)	(0.033)	(0.341)	(0.359)
Constant†	0.749***	0.714***	0.346*	0.457***	0.502***	0.489***	0.214***	0.095**	5.581***	6.516***
	(0.046)	(0.055)	(0.180)	(0.144)	(0.052)	(0.061)	(0.065)	(0.046)	(0.505)	(0.617)
Covariates‡	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	1,710	1,513	1,216	1,039	1,710	1,513	820	712	801	702
Mean of outcome§	0.506	0.699	0.556	0.739	0.375	0.428	0.265	0.140	5.759	6.576
R-squared	0.053	0.042	0.053	0.092	0.077	0.075	0.302	0.327	0.092	0.130

Notes: This tables summarizes estimation results from OLS regressions for five outcome variables, in two different subsamples (low versus high physician's type-specific density). SHI stands for statutory insurance. PHI stands for private insurance. Robust standard errors are reported in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. † The base group (=constant) is equal to Yilmaz, SHI. ‡ Covariates comprise physician's type fixed-effects, and city fixed-effects. § Refers to base group.

Table A.7: Estimation results by type of physician

	All	Dentists	Ophthalm-ologists	Dermat-ologists	Orthopedists
	(1)	(2)	(3)	(4)	(5)
Schmidt, PHI	0.129***	0.054	0.205***	0.163**	0.211***
	(0.025)	(0.037)	(0.059)	(0.075)	(0.058)
Schmidt, SHI	0.020	0.029	0.057	−0.074	0.019
	(0.025)	(0.036)	(0.055)	(0.071)	(0.057)
Yilmaz, PHI	0.134***	0.040	0.254***	0.205***	0.184***
	(0.025)	(0.036)	(0.057)	(0.076)	(0.059)
Constant†	0.432***	0.507***	0.485***	0.365***	0.332***
	(0.033)	(0.062)	(0.070)	(0.081)	(0.066)
Covariates‡	Yes	Yes	Yes	Yes	Yes
Number of observations	3,224	1,552	638	405	629
Mean of outcome§	0.401	0.496	0.333	0.325	0.281
R-squared	0.042	0.074	0.171	0.217	0.173

Notes: This tables summarizes estimation results from OLS regressions for a binary outcome variable indicating the offer of an appointment in five different (sub)samples. We distinguish between the full sample (comprising all types), and sub-samples for dentists, ophthalmologists, dermatologists, and orthopedists. SHI stands for statutory insurance. PHI stands for private insurance. Robust standard errors are reported in parenthesis. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. † The base group (=constant) is equal to Yilmaz, SHI. ‡ Covariates are city fixed-effects. § Refers to base group.