

NBER WORKING PAPER SERIES

CREAM SKIMMING BY HEALTH CARE PROVIDERS AND
INEQUALITY IN HEALTH CARE ACCESS:
EVIDENCE FROM A RANDOMIZED FIELD EXPERIMENT

Anna Werbeck
Ansgar Wübker
Nicolas R. Ziebarth

Working Paper 28809
<http://www.nber.org/papers/w28809>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02139
May 2021

We thank Sebastian Bauhoff, Kevin Devereux, Peter Eibich, Jenke De Keyser, Brigitte Dormont and Harald Tauchmann for excellent discussions of this work. We also thank Tigran Aydinyan, Sergej Bechtoldt, Katharina Blankart, Amitabh Chandra, Alberto Holly, Daniel Herrera, Tom Hoe, Toshiaki Iizuka, Nima Jouchaghani, Hendrik Jürges, Helene Könnecke, Christoph Kronenberg, Isabelle Maripier, Giuseppe Moscelli, Julien Mousquès, Catia Nicodemo, Cheti Nicoletti, Simon Reif, Hendrik Schmitz, Jonas Schreyögg, Hannes Schwandt, Erin Strumpf, David Slusky, Christian Waibel, Chris Whaley, Daniel Wiesen, Lilia Zhurakovska, Peter Zweifel and participants at the 2020 Annual Meeting of the Swiss Society of Health Economists (sggö), the 2019 Annual Meeting of the German Society of Health Economics (dggö), the 2019 Health Economics Workshop in Stralsund, 2019 American-European Health Economics Study Group, the 2019 ASHEcon meetings, the 2019 Annual Congress of the European Economic Association, the 2019 IRDES-DAUPHINE Workshop on Applied Health Economics and Policy Evaluation, the 2019 RGS conference, the 2018 Workshop of the dggö committee „Gesundheitsökonometrie“ (“health econometrics”) for helpful suggestions and comments. Special thanks go to Alexander Hearing. We also thank Paul Pangburn for excellent research assistance. We thank WIdO and an anonymous private insurer for supplemental insurance data. The research reported in this paper is not the result of a for-pay consulting relationship. Our employers do not have a financial interest in the topic of the paper which might constitute a conflict of interest. The field experiment was carried out at RWI Essen, a German publicly funded research institute, which does not have IRB approvals. All ethics guidelines by the German Science Foundation were followed; the guidelines for good scientific practice are here http://www.dfg.de/en/research_funding/principles_dfg_funding/good_scientific_practice/. The standardized protocols minimized the time to inquire about appointments. No actual appointment was made. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2021 by Anna Werbeck, Ansgar Wübker, and Nicolas R. Ziebarth. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Cream Skimming by Health Care Providers and Inequality in Health Care Access: Evidence from a Randomized Field Experiment

Anna Werbeck, Ansgar Wübker, and Nicolas R. Ziebarth

NBER Working Paper No. 28809

May 2021

JEL No. I11,I14,I18

ABSTRACT

Using a randomized field experiment, we show that health care specialists cream-skim patients by their expected profitability. In the German two-tier system, outpatient reimbursement rates for both public and private insurance are centrally determined but are significantly higher for the privately insured. In our field experiment, following a standardized protocol, the same hypothetical patient called 991 private practices in 36 German counties to schedule appointments for allergy tests, hearing tests and gastroscopies. Practices were 4% more likely to offer an appointment to the privately insured. Conditional on being offered an appointment, wait times for the publicly insured were twice as long than for the privately insured. We also find smaller access differences when reimbursement rate differences are smaller. Our findings show that structural differences in reimbursement rates lead to structural differences in health care access.

Anna Werbeck
RWI Essen
Hohenzollernstraße 1-3
Essen 45128
anna.werbeck@rwi-essen.de

Ansgar Wübker
RWI Essen
ansgar.wuebker@rwi-essen.de

Nicolas R. Ziebarth
Department of Policy Analysis and Management
Cornell University
2218 MVR
Ithaca, NY 14853
and NBER
nrz2@cornell.edu

1. Introduction

Access barriers to health care are a major performance indicator in comparative health care system analysis (Siciliani and Hurst 2005; Siciliani and Verzulli 2009; Jones et al. 2011, Viberg et al. 2013). The nonpartisan *Commonwealth Fund* uses wait times as the main measure of health care access in their “Timeliness to Care” category, where the United States ranks 9th among 11 countries in the 2017 survey (Commonwealth Fund 2017). At the same time, wait times have long been cited by critics as proof of mediocre outcomes of single-payer systems (cf. Mackillop et al. 1995). Indeed, wait times for specialists are significantly longer in Canada as compared to the largely private system in the United States. In Canada, 30% of patients have to wait more than 2 months for a specialist appointment as compared to just 6% in the United States (Commonwealth Fund 2017).

Another major performance indicator to rate health care systems is *equity* in access to health care (e.g. van Doorslaer et al. 2000)—a dimension on which the United States has consistently ranked last among the 11 OECD countries benchmarked by the *Commonwealth Fund* (2014, 2017). In the U.S., thousands of private managed care insurers individually negotiate reimbursement rates with networks of providers. Furthermore, the public Medicaid system for the poor pays significantly lower rates than private insurers or the single-payer Medicare system for the elderly (CMS 2018a). Critics of this fragmented private-public U.S. system have pointed out large structural inequalities, even among those who have insurance (Sommers et al. 2017). One consequence of a system with major differences in reimbursement rates could be that providers cream-skim and discriminate against those with structurally lower reimbursement rates which tend to be the poor, the sick as well as minority groups (Reinhardt 2011). However, although plenty of anecdotal and descriptive evidence exists, it is difficult to show in a causal framework that health care providers cherry-pick the privately insured *because* they are more profitable.

This paper uses a randomized field experiment in a well-suited institutional private-public payer setting to show that health care specialists cream-skim the more profitable privately insured patients. Germany has a multi-payer two-tier system where the majority of the population is mandatorily insured under the public system in one of the 110 non-profit public “sickness funds” (Schmitz and Ziebarth 2017). In the public system, provider reimbursement rates are centrally negotiated and do not vary across sickness funds. Moreover, cost-sharing is standardized and invariant across sickness funds, while provider networks are non-existent and enrollees can freely choose their provider (Bauhoff 2012, Bünnings et al. 2018). The situation is similar for the 9 million privately insured residents: reimbursement rates are uniform across the 44 private insurers and provider networks do not exist; insurers mostly process claims (Atal et al. 2019). However, in the outpatient sector, reimbursement rates for the privately insured are on average more than twice as high than for publicly insured (Walendzik et al. 2008). This institutional setup is well-suited for our study. No other country has a two-tier public-private health care system without provider networks and with reimbursement rates that (i) structurally vary between the two systems (ii) but are otherwise identical across plans within each system.

In our field experiment, we selected a total of 36 representative counties (both urban and rural) and called a total of 991 outpatient specialists to ask for appointments for elective medical treatments. One single test person called each practice *twice*, once as a fictitious privately insured new patient and once as a fictitious publicly insured new patient, randomizing the order of the insurance status between the two calls. In other words, the same test person called each private outpatient practice twice following the exact same protocol, thereby ensuring balanced covariates by construction. One main advantage of calling each practice twice is that it allows the researcher to exploit within-practice variation when analyzing wait times, conditional on appointments being made. In other words, even

conditional on appointments actually being offered, the randomized insurance status is exogenous. The research design allows us to carry out straightforward statistical tests to assess whether extensive and intensive access barriers to health care differ significantly by insurance status.

Our findings show that access to the health care system differs significantly between the privately and publicly insured, both on the extensive and the intensive margin. The likelihood that practices offer an appointment is a highly significant 4% larger for privately insured patients. Moreover, conditional on being offered an appointment, the wait times for publicly insured patients are more than twice as long, and on average 13 weekdays longer. We also find that access discrimination is not as pronounced when the reimbursement rate differential is smaller.

This paper makes important contributions to the literature. Although the literature on physician behavior and treatment styles is rich and has a long tradition in economics (e.g., Clemens and Gottlieb 2014, also see Section 2), the causal effects literature on how providers discriminate against less profitable patients is less diverse. We contribute to a better understanding of the role of varying reimbursement rates in determining equitable access to the health care system for disadvantaged population groups. For example, for the United States, Cooper et al. (2018) document that reimbursement rates just among the privately insured could vary by a factor of 10 within cities and by more than 20 across the United States of America.

Natural experiment studies closely related to this research have investigated whether the Medicaid Fee bump¹ of 2013 and 2014 has increased health care access for low-income populations in the U.S. Although studies outside economics solely investigate statistical associations (Polsky et al. 2015; Saloner et al. 2015; Candon et al. 2018), the evidence by two economic causal effect studies

¹ The Medicaid fee bump is a provision of the Affordable Care Act that mandates states to increase Medicaid payments to match Medicare rates for primary care visits for 2013 and 2014.

suggest that the Medicaid Fee bump likely did increase access among the poor (Alexander and Schnell 2017; Maclean et al. 2018).

Our study is one of the first real-world studies that leverages a large randomized field experiment across an entire country—the fourth largest economy in the world—to test if the insurance status causally affects health care access because of a higher expected profitability of the privately insured. For that purpose, we focus on one of the few countries worldwide with co-existing private-public insurance, a system without confounding factors such as managed care or restricted provider access. Moreover, we know the exact reimbursement rates for the treatments requested, and we rely on actual claims data by ICD-10 diagnoses from one of the largest public and one of the largest private insurers.

In our experiment, one single test person called almost one thousand providers twice where we randomized the order of the insurance status of the caller. Unlike the few existing studies outside the field of economics, to minimize selection concerns, our caller routinely informs providers about the insurance status and inquires wait times without further framing. Moreover, we focus on non-urgent routine specialist visits, which have external validity for everyday interactions between patients and providers in a multi-payer system with a public-private mix of insurers. We believe that our findings have external validity for similar systems, such as the U.S. system where healthier, wealthier and racially less diverse populations are typically covered by better paying private insurance, whereas the public Medicaid system pays doctors much lower rates and covers sicker and poorer populations.

Our findings yield important insights into the driving forces of inequity in health care access (cf. Chen et al. 2019). They suggest that uniform reimbursement rates (or reimbursements rates that are higher for disadvantaged population groups) could help mitigate inequality in health care access and

align economic incentives with medical needs and priorities. At the same time, we do not dispute that differences in reimbursement rates can be the outcome of an efficient system if patients have the choice between differently priced plans and make informed decisions (cf. Handel and Kolstad 2015).²

2. Previous literature

This paper relates to various literature strands in economics. However, while many descriptive papers on socio-economic differences in health care access exist, the causal effects literature on discrimination in the health care sector is much thinner.

The economics literature has a long tradition of investigating theoretically and empirically the role of physicians as (imperfect) agents of their patients, see McGuire (2000) for an excellent overview. In addition, economists have investigated how physician behavior and productivity changes in response to the reimbursement method, in outpatient as well as in inpatient settings (Ellis and McGuire 1986; Nicholson et al. 2008). Baker and Royalty (2000) show in the U.S. context that expanded Medicaid eligibility increased access to physician services. Similarly, Decker (2009) finds that cuts in Medicaid physician fees reduced the number of visits for Medicaid patients compared to privately insured patients. For Germany, Schmitz (2013) shows that newly introduced practice-level budgets for the publicly insured reduced the likelihood of follow-up visits. Reif et al. (2013) use administrative data on newborns and several discontinuities in birth weight thresholds that trigger higher reimbursement rates to test for whether treatments quantity and quality differs. While they find evidence for weight manipulation, they conclude that there is not much evidence for causal revenue-related changes in treatment quantity and quality at these margins in Germany.

² However, the large majority of Germans *cannot* choose between the private and public system; the institutional rules separate the public from the private insurance market (see Section 3).

For ethical reasons, real-world field experiments are almost impossible to implement to study actual treatment behavior, which is why researchers have conducted audit studies (Bauhoff 2012) or investigated hypothetical physician behavior in the lab; see Brosig-Koch et al. (2017) and Kesternich et al. (2015) for lab experiments in Germany. For example, Kesternich et al. (2015) find among 266 medical students that professional norms shift more weight to patients and away from own profit and efficiency concerns. In one of the few real-world causal effects studies leveraging relative price changes in the Medicare outpatient market, Clemens and Gottlieb (2014) demonstrate that higher relative reimbursement rates increase treatment volume, especially for elective procedures.

Absent price variation in single-payer markets, implicit rationing of medical care through wait times is another popular topic of inquiry for economists (e.g. Lindsay and Feigenbaum 1984). Cullis et al. (2000) provide a comprehensive overview of the topic, and Siciliani (2006), Gravelle and Siciliani (2008), and Felder (2008) theoretical analyses. The link between wait times and socio-economic status has also drawn researchers' interest. For example, Monstad et al. (2014) find a negative statistical correlation between income and wait times as well as education and wait times in Norway. Laudicella et al. (2012) show that the same correlations exist in England and that they hold up over the entire wait time distribution.

The impact of insurance status on wait times is a highly policy relevant topic in countries with co-existing insurance systems that pay providers differently, such as the United States, Switzerland and Germany. Hence, several papers have studied the link between insurance status and wait times of patients (Roll et al. 2012; Sundmacher and Kopetsch 2013; Ramos et al. 2018). All of them find that patients whose insurer pays lower rates have to wait longer for an appointment. However, because enrollment in "lower rate" insurers such as U.S. Medicaid is correlated with specific socio-demographics as well as Managed Care elements such as gatekeeping or capitation, it remains challenging to identify

causal effects of insurance status on discrimination through providers. Similar arguments hold for the case of Germany. For example, Jürges (2009) uses rich SOEP survey data and find in multivariate regression models—conditioning on self-reported health—that the privately insured are less likely to contact a physician but more likely to have a follow-up treatment after the first contact. Hullegie, Klein and Salm (2012) use the same data and exploit the discontinuity to opt out of the public system to conclude that the private insurance reduces doctor visits.

The paper that is closest to ours is Lungen et al. (2008). In April and May of 2006, Lungen et al. (2008) called 189 specialist practices in one region in Germany (in the neighboring cities of Cologne, Bonn and Leverkusen). After excluding a third of all practices, they find that practices offered appointments to the privately insured three times faster than the publicly insured. One objective of this study is to test whether these findings hold up when using a much larger representative sample size, an improved methodology, and recent data. The main differences to our study are:

First, Lungen et al. (2008) only called practices once and thus cannot exploit within-practice variation using practice fixed effects; moreover, we show that our randomization was successful and the covariates balanced. Second, in their models, Lungen et al. (2008) only control for the specialty, whereas we control for county and practice fixed effects as well as the time of the day, the day of the week and the week of the year. Third, Lungen et al. (2008) study wait times using a GLM model while we study both, the important extensive as well as the intensive margin. Fourth, whereas Lungen et al. (2008) conduct their experiment about 15 years ago, our experiment happened between 2017 and 2018. It thus considers the current institutional background of the German health care system. Fifth, we called nearly 1000 practices over the course of an entire year in 36 countries, which we chose to ensure representativeness. Sixth, the sample size allows us to study differences by population and physician density, East and West Germany, the share of PHI insured in the market and the average

income in the county. Finally, we use secondary claims data as well as calculated reimbursement rates from the SHI and PHI fee-for-service schedules to trace wait time differences back to actual differences in reimbursement rates. This is important to exclude that other structural differences between the two groups are driving the effect.³

In addition to Lungen et al. (2008), four studies called providers at least twice with the insurance status randomized. However, two of them focus on the United States (Asplin et al. 2005, Skaggs et al. 2001), where exact reimbursement rates are largely unknown and the utilization management of managed care organizations and provider networks likely act as confounders. Kuchinke et al. (2009) use the German setting but called 500 acute care *hospitals* in Germany, where reimbursement rates do not vary between public and private insurance.⁵ Fourth, Heinrich et al. (2018) called 163 specialists in one region and *evaluated a 2015 reform* that intended to reduce wait times for the publicly insured in Germany.

In our setting, a test person deliberately called each practice twice following a standardized protocol, where the (order of the) insurance status of the caller was randomized. Moreover, the *same person* called all private specialist practices and indicated the insurance status when trying to schedule a non-urgent medical examination. In Germany, telephone calls are the most common and most natural approach to schedule appointments. We use a contemporaneous setting over the course of one year

³ Two field experiments estimate discrimination based on socioeconomic status in Canada (Olah et al. 2013) and Austria (Angerer et al. 2019). The former study called 375 family practices in Toronto. They find that mentioning being an employee of a major bank vs. a welfare recipient increases the likelihood of getting an appointment. The latter study requested 1,310 appointments by email. They find that closing with “Dr.” increases the likelihood of an appointment. In another study outside of economics, Resneck et al. (2006) called 631 dermatologists once and randomized insurance status. However, the practices were framed and knew that they would participate in a study and the main outcome was whether they would *accept* Medicaid, Medicare and privately insured, in general. Finally, descriptive survey data from the National Ambulatory Medical Care Survey show that office-based physicians in the United States are more likely to accept new patients from private insurance compared to new patients from Medicare or Medicaid (cf. Decker 2012 or Holgash and Heberlein 2019).

⁵ In a spin-off paper, Schwierz et al. (2011) investigate effect heterogeneity and differentiate the findings by the financial soundness of the hospital.

between 2017 and 2018. We called almost one thousand practices located in a representative set of 36 German counties—more than any other study, which allows us to carry out a detailed subgroup analysis. We focus on elective care among outpatient specialists as non-urgent settings are those where most patient-provider contacts occur. These treatments reflect the regular day-to-day access barriers to health care much better than studying medical emergencies. In addition, we know and calculate the centrally negotiated differences in reimbursement rates for the treatments requested and also investigate actual claims data for one of the biggest public and one of the biggest private insurers in Germany. That circumvents an inherent problem of U.S. studies where single providers and insurers negotiate thousands of largely unobserved reimbursement rates along with complex differences in insurer-provider networks under managed care, which also include cost-sharing elements and various utilization management tools.

3. The German Health Care System

Germany has a two-tier health insurance system with a co-existing multi-payer public and an individual private market. Ninety percent of the population are covered by the public system and one of the 110 non-profit sickness funds (GKV Spitzenverband 2018). They pay income-dependent contribution rates for a standardized benefit package with very little cost-sharing. For historical reasons, selected population subgroups have the right to leave the public system permanently and fully insure their health risks on an individual long-term health insurance market with relatively little regulation. Applicants can choose between thousands of plans but are also experience-rated when signing their first individual private contract (in subsequent years, premiums are community rated). Schmitz and Ziebarth (2017), Pilny et al. (2017), and Bünnings et al. (2018) provide more details on the overall structure of the German health insurance market. Atal et al. (2019) provide additional specific

details of the private market. Note that 64 million Germans, or 77% of the total German population, are *mandatorily* insured with the public scheme (BMG 2019). Those people do *not* have the choice between public or private insurance.⁷

Table A1 in the Appendix uses representative data from the German Socio-Economic Panel Study to compare mean characteristics of the publicly and privately insured in Germany. The first column shows sociodemographic averages for the publicly insured and the second column shows sociodemographic averages for the privately insured. The last three columns further differentiate by the four population subgroups that can be privately insured (civil servants, high income, self-employed, non-employed). As seen, the privately insured—not just on average but also in all four subgroups separately—have significantly higher gross wages (€4,708 vs. €2,403) and significantly higher post-tax post-transfer household incomes per person (€40,031 vs. €23,228). They are on average more than 3 years older, are 10 percentage points less likely to be smokers, have lower BMIs and report fewer physical and mental health limitations. They are also less likely to be hospitalized and have fewer hospital days per year. Interestingly, their number of outpatient visits is identical to those that are publicly insured.

Reimbursement Rates in Statutory Health Insurance (SHI)

In the outpatient SHI sector, primary care physicians and specialists are members of and sign contracts with the state-level “Regional Association of Statutory Health Insurance Physicians”, ASHIP

⁷ Since 2007, SHI policyholders have the option to enroll in so called “optional tariffs” (*Wahltarife*). One of those is called *Kostenerstattungstarif* and allows outpatient providers to charge the same higher rates as for PHI policyholders, the difference of which has to be paid out of pocket by enrollees and is only partially reimbursed by insurers. Due to this unappealing feature and also because these possibilities are not widely known in the general population, the demand for the *Kostenerstattungstarif* is very low and only an estimated 0.5% of all SHI enrollees chose it in 2010 according to the SOEP (2016).

(*Kassenärztliche Vereinigungen*), see KBV (2018a). There are 17 ASHIPs, who are responsible for the provision of health care services in their region. These ASHIPs all have contracts with the 110 sickness funds who pay out a “total reimbursement sum” (*Gesamtvergütung*) to each of these 17 ASHIPs who, in turn, reimburse their member physicians on a quarterly basis.

In SHI, the so-called “Unified Assessment Scale” (*Einheitlicher Bewertungsmaßstab, EBM*) lists services that the SHI benefit package covers. The existence of the EBM is stipulated by the German Social Insurance Law (§ 87f. SGB V, KBV 2018b). The EBM assigns a point value for each health care service, similar to the Relative Value Units (RVU) to outpatient providers in Medicare in the U.S. (CMS 2018a). The relative point values intend to represent the relevant use of resources for each service to provide an adequate compensation.⁸

Similar to Medicare, by defining values per point annually, the point values are then converted into monetary reimbursement amounts. For example, in 2018, the value per point was 10.654 euro cents (BMG 2018).⁹ For a colonoscopy for preventive reasons, including visits to prepare and inform the patient, the EBM lists 1945 points under “fee position” (*Gebührenordnungsposition*) 01741 (KBV 2018b). Consequently, the total basic compensation for such a colonoscopy would be €207.23.¹⁰ In comparison, for the state-level Medicaid insurance for low-income populations in the United States,

⁸ In SHI, when physicians provide more services than allocated by the “standardized service volume” (*Regelleistungsvolumen*), which is defined by the sum of last quarter’s services and the average in the specialist group, the point value can decrease (§87b SGB V). However, since 2012, these budget caps, set by the sickness funds in cooperation with the ASHIPs are optional, see Simon (2017). For example, North Rhine still imposes budget caps (KVNO 2018).

⁹ Geographic adjustment factors take differences in regional living costs into account.

¹⁰ Interestingly, the reimbursement rates for colonoscopy in the U.S. under Medicare are similar. Under CPT code 45380 “Colonoscopy and biopsy”, the Medicare fee schedule lists a reimbursement of \$212.70 (CMS 2018b).

Halpern et al. (2014) report reimbursement rates between \$83.94 in New York and \$598.20 in Alaska for a colonoscopy.

Reimbursement Rates in Private Health Insurance (PHI)

In PHI, the physician has a private contract with the patient. Patients have to pay providers first (after receiving an invoice), and then submit their claim to the insurer to get reimbursed. In PHI, the “Fee Schedule for Physicians” (*Gebührenordnung für Ärzte, GOÄ*) lists all reimbursable services along with their baseline prices. As with SHI, each medical service has a specific number and point value; the latter expresses the relative resource utilization for the treatment. Point values are multiplied with a fixed value of 5.82873 euro cents to obtain the baseline reimbursement rate.

Depending on the complexity of the treatment and the time spent on its provision, the physician has the freedom to multiply the baseline rate with “adjustment factors” between 1.15 for laboratory services and up to 3.5 for personal services (§5 II-IV GOÄ).¹¹ Also, the physician can perform and charge treatments not listed in the GOÄ, using prices of similar treatments as a reference (Simon 2017). Overall, the GOÄ is a classic fee-for-service schedule without any budget caps or cost containment elements. For example, a standard colonoscopy is listed as number 687 with 1500 points and a baseline value of €87.43 (GOÄ 2018).

Comparison of the SHI and PHI Reimbursement

A direct comparison of the SHI and the PHI reimbursement scheme is not straightforward as the treatments and services in each schedule usually do not exactly correspond. Walenzik et al. (2008)

¹¹ If the adjustment factor is larger than 2.3, the calculation must include a justification for the multiplier chosen regarding the degree of difficulty and time required for the treatment.

analyze and compare differences in the reimbursement amounts for the same treatments under SHI and PHI using data from the largest German sickness fund. They find that providers charge 2.28 times higher reimbursement rates for privately insured patients.

One strength of our institutional setting is that price schedules are centrally set for both the SHI and PHI, and not individually negotiated between insurers and providers. In addition to being able to look up the exact list prices for the treatments of our experiment, we obtained access to two claims databases, one by the largest public and one by the largest private insurer. This allows us to compare the average claim submitted to the insurer by ICD-code. Table 1 shows a comparison of what providers typically charge for the medical examinations inquired in our experiment (KBV 2017, 2018b, GOÄ 2018).

The first two columns of Table 1 show the points and euro values for publicly insured patients. Column (3) lists actual claims by ICD-10 code submitted to one of the biggest public insurers with a market share of roughly 35 percent (*Allgemeine Ortskrankenkassen*). It indicates the total claim amounts when the diagnosis contained ICD-10 code T78 (“allergies”) or ICD-10 codes K29 (“Gastritis and duodenitis”), K30 (“Functional dyspepsia”) or K31 (“Other diseases of stomach and duodenum, WIdO (2020).

[Insert Table 1 about here]

Columns (4) and (5) yield the points and euro values for the privately insured and column (6) lists actual claim amounts by the same ICD-10 codes as above, but submitted to one of the largest *private* insurers (Karlsson et al. 2016).¹² As seen, list prices for allergy tests are more than three times

¹² Note that managed care basically does not exist in Germany; private insurance essentially represents a purely financial contract (cf. Atal et al.2020).

higher for privately insured (€49 vs. €184); actual submitted claims that contain ICD-10 codes T78 (“allergies”) are about twice as high (€110 vs. €201). The absolute and relative price differences for hearing tests are much smaller but rates for the privately insured are still higher (€16 vs. €23). For upper gastrointestinal endoscopies, like for allergy treatments, list prices for the privately insured are more than two times higher (€89 vs. €204) and submitted claims for specific diagnoses K29-31 are also significantly higher (€172 vs. €245).

Finally, the Federal Statistical Office provides detailed statistics about the net revenue (revenue after costs) of outpatient practices by specialty and type of practice. Accordingly, the net revenue per specialist owning a practice was €183K for ear-nose-throat (ENT) doctors, €206K for internal medicine and €225K for dermatologists (Destatis 2018c).

4. The Experiment

Selection of Counties

Before sampling outpatient providers, we first selected a set of counties that are jointly approximately representative for Germany. We considered the following three indicators: household income per capita, area in square kilometers, and the population (BBSR 2018; Destatis 2018a, b). Appendix B describes in detail how we selected the counties.

[Insert Figure 1 about here]

Figure 1 shows Germany with its 401 counties, where the dark gray-shaded counties are part of the field experiment. As seen, the geographic distribution of all 36 counties is relatively even across all 16 German states as well as between East and West Germany. Comparing the monthly household income per capita of the 36 counties to the monthly household income for the whole of Germany, we

only find minor differences (€1,723 vs. €1,753). Also, the physician density per 100,000 population is almost identical when comparing the 36 counties to Germany as a whole (174 vs. 168 physicians per 100,000 population, see Versorgungsatlas 2018).

Sampling of Outpatient Specialists and Treatments

Next, for these 36 counties, we randomly sampled outpatient specialists using Google Maps along with the websites of the three major German telephone books to identify currently operating outpatient specialists in each of the 36 counties; we used “The Telephone Book”, “Yellow Pages” and “The Local” (Das Telefonbuch 2018).

In a pre-test, we called specialists anonymously and scheduled appointments for six different non-urgent medical examinations in the cities of Berlin, Cologne, Bonn, Leverkusen, Hamburg and Munich. The examinations were an allergy test, a hearing test, an eye examination, a gastroscopy, a magnet-resonance-therapy of the right knee, and a pulmonary function test.

After this pre-test, in the remaining 30 counties, using the exact same protocol than in the pre-test, we called gastroenterologists, otorhinolaryngologists, and allergists to schedule appointments for the following three examinations: (a) an upper gastrointestinal endoscopy, (b) an audiometry (hearing test), and (c) an allergy test. We chose these three (out of six) examinations because they are the most popular, non-urgent routine examinations and are relatively easy to schedule.¹³

¹³ Making appointments and eliciting waiting times for the other three examinations also proved less feasible in the pre-test. For example, we received more reasons for practice exclusions and queries about the experiment. Nonetheless, in principle, the other three examinations would also have been suitable for the study. Due to limited resources for larger samples, we could not generate the sample sizes necessary for all six examinations; hence, we prioritized and focused on the three specialties. Note, in our baseline sample, we also include pre-test data. In a robustness check, we test whether the results are robust to the exclusion of pre-test data (see Section 7).

Study Design

In total, we called 991 private practices to schedule appointments. The same test person (the “caller”) made the calls over the course of one calendar year, between April 6, 2017 and May 3, 2018. Importantly, the test person called each practice *twice* and clearly indicated the insurance status of the fictional patient.¹⁴ We randomized whether the caller would pretend to be privately or publicly insured. Moreover, we made the two calls in time intervals of at least two weeks to not raise any suspicion about being part of an audit study. Specifically, we randomized the (order of the) insurance status of the caller, conditional on (i) day of week, (ii) time of day and (iii) week of the year. However, we did not randomize (i) to (iii) which is why we control for these variables in our regression models, and why balance tables would only show variation in (i) to (iii) but in none of the other variables.¹⁵

During each call, we followed a pre-determined standardized protocol on how to start and end the call and what answers to give in response to the most frequently asked questions (see Appendix C). We made all calls between Monday and Friday during the regular office hours of each practice.¹⁶ During the call, the caller mentioned that she had a referral by her Primary Care Physician (PCP). When asked

¹⁴ When asked about the name of the insurer, we provided a real name. Because reimbursement rates are centrally determined for publicly and privately insured (Section 3) and not individually negotiated between insurers and providers, the actual insurer is not crucial in the German setting. Note that we did not disclose to the practice manager or receptionist who took the call that they were part of an audit study, as that information would bias the results of the research project. Riach and Rich (2004) argue that lack of veracity in the context of discrimination studies may justify deception of participants. Similarly, Levitt and List (2009) argue that there are valid arguments for not making informed consent the rule in a field experimental context. This argument holds particularly if informed consent directly interferes with the ability to conduct the research (Levitt and List 2007).

¹⁵ Imagine we would not randomize and always pretend to be publicly insured on the first of the two calls. Then we could not disentangle insurance status effects from week-of-year seasonal effects. Depending on the specific setup, such a nonrandomized design may also confound time-of-day and day-of-week effects with treatment effects. Obviously, supply and demand side factors vary systematically by day of week, time of day and week of year. We thank an attentive referee for this remark.

¹⁶ If voicemail indicated special office hours, the follow-up calls were made during these special hours. When nobody answered the phone, the practice was flagged as “not available” after three unsuccessful attempts. When the line was busy in one of these three attempts, the maximal number of attempts was raised to six.

for the name of the PCP, the caller gave a fictional name and indicated that the practice was in her hometown. Note that, unlike previous studies such as Kuchinke et al (2009), we routinely disclosed the insurance status of the caller.¹⁷ This is very common in the German outpatient setting and does not raise suspicion (Spiegel 2015). Moreover, it facilitates the empirical analysis as insurance status disclosure is not endogenous. Finally, the caller ended all calls without fixing the suggested appointment to not occupy a slot that could be used for a real treatment. Also, recall that all requests were for elective non-urgent treatments.

In total, we called 991 unique private practices in the 36 German counties displayed in Figure 1. Figure A1 in the Appendix shows the distribution of the contacted practices across the 36 counties. The number of contacted practices varies between 1 in two very small and low populated counties and 126 in one big German city. The mean number of practices contacted was 26 per county. In most counties, all three specialists were available.

5. Data

Sample Selection

First, we exclude practices from our study for the following reasons: (i) the specialist is not active anymore (19, 1.9%), (ii) the practice only offers treatments for privately insured patients¹⁸ (43, 4.3%),

¹⁷ In Kuchinke et al (2009), who study hospitals, only 25% inquired about the insurance status. However, this is not surprising as hospital reimbursement rates do not vary between publicly and privately insured in Germany. Only in case of selected special services like treatments by leading senior physicians ("Chefarztbehandlung") or for special service quality amenities like private rooms, hospital revenues differ structurally between publicly and privately insured.

¹⁸ Practices have the option to entirely opt out of treating publicly insured patients and declaring themselves an exclusive practice for privately insured only and people who pay entirely out-of-pocket. These practices, however, are then not eligible to charge sickness funds, even when demand from private patients is low. We do not consider these practices as relevant to the experiment.

and (iii) other reasons¹⁹ (62, 6.3%). These reasons reduce the number of unique practices in our study by 124 from 991 to 867.

Second, there were additional reasons why practices were unresponsive and we could not make appointments; e.g., the (a) practice was closed for at least one week, for example during vacations; (b) the practice did not make fixed appointments; (c) the practice did not accept new patients temporarily or (d) the practice was not reachable due to several unsuccessful attempts. Some of these practices were unresponsive in one wave (110 in wave 1 and 94 in wave 2) and others were unresponsive in both waves (101 practices). With 766 practices, we could successfully inquire about an appointment at least once. In other words, for all eligible practices that are not entirely excluded from our sample due to reasons (i) to (iii) above, we either tried to make an appointment during the first call under insurance status A, during the second call under insurance status B, or during both calls. We call this unbalanced sample “Sample A;” it has 1,328 observations from 766 unique and successfully contacted practices. Figure A4 shows a sample selection chart and Table 2 shows the descriptive statistic for this full sample.

By contrast, our “Sample B” only includes caller-appointment observations where the practice offered an appointment to *both* fictional patients, the publicly and the privately insured. This sample is balanced, includes 502 unique private practices, and 1,004 caller-appointment observations.

Main Outcome Variables

We generate two main outcome variables, both of which measure access to the health care system. The first variable is binary and called *apptm*. It indicates whether the successfully contacted

¹⁹ E.g. practices for children only and misleading telephone numbers.

practice was willing to schedule an appointment with the fictional patient. As seen in Table 2, in 92% of all cases, the practice was willing to schedule an appointment.

The second variable is continuous and called *dayswait*. It counts the number of workdays (which equal weekdays) from the calling date to the offered appointment.²¹ It has only valid values when the practice offered an appointment. Figure A2 shows the distribution of *dayswait* and Table 2 shows the summary statistic. As seen, the minimum wait time is an immediate appointment, when patients could be seen on the same day. The maximum wait time is 171 weekdays and the average wait time is 19 weekdays (almost 4 weeks). Figure A2 shows a left-skewed distribution with a long right tail.

Main Control Variables

The main variable of interest is *Privately Insured*. Even in the unbalanced Sample A with 1,328 caller-appointment observations, this variable is almost perfectly balanced with 50.1% of all observations representing a privately insured fictional patient (Table 2).²²

Other important control variables indicate the day-of-the-week, the exact calendar date, the time of the day when we made the call, whether the randomized insurance status was privately or publicly insured during the first call, and the specialty of the practice. Recall that we randomize the

²¹ We excluded weekends (Saturday, Sunday) as well as public holidays. In a robustness check, we also excluded weekdays between a public holiday and weekends (*Brückentage*) as many Germans take vacation days during these days to have an extended weekend off. We call this variable *dayswait II* (see Table 1).

²² We report separate descriptive statistics for practices that only responded once vs. practices that responded twice in Table A2. As seen, practices that successfully responded only once i) have a slightly higher average waiting time than those that responded twice and ii) show differences in other variables and county-level controls, for example, a higher share of privately insured and East German practices. To account for that imbalance, we conduct additional robustness checks in Section 7.

insurance status conditional on calling a specific practice on a specific day, which implies that controlling for these variables is relevant.

County-Level Control Variables

The final panel in Table 2 lists the county-level control variables. These have been provided by the Federal Statistical Office (Destatis 2018a, b) and by the Federal Institute for Construction, Urban and Space Research (*Bundesinstitut für Bau-, Stadt- und Raumforschung*), see BBSR (2018). As seen, the average net income in the 36 counties is €1,789 and the average unemployment rate is 7.9%.

6. Statistical Methods

Most important for causal inference in this setting is the fact that we set up a field experiment, where a test person called each specialist practice twice, where we randomized the order of the insurance status. Calling each practice twice and randomizing the insurance status guarantees balanced covariates by design (except for day-of-week, time-of-day and week-of-year, see above). Moreover, the design allows us to exploit within-practice variation when analyzing the intensive margin, conditional on appointments being offered. Because the same test person called all 991 practices and strictly followed a pre-determined protocol, simple descriptive statistics and nonparametric bivariate tests should yield first reliable evidence about access differences between the two insurance groups.

As our main statistical approaches, we run OLS and count data regression models which routinely control for the calendar date, the day-of-the-week and the time during the day of the call—in addition to practice-level and county-level controls. Our first model uses the unbalanced Sample A:

$$a_{ip} = \alpha + \beta PHI_i + X_p' \tau + Z_c' \theta + \gamma DOW_{ip} + \zeta TOD_{ip} + \delta_t + \rho_p + e_{ip} \quad (1)$$

where a_{ip} stands for our first outcome variable *apptm*, which is binary and indicates whether practice p offered the fictional patient i an appointment or not. The main variable of interest is PHI_i and indicates whether the caller indicated a public or private insurance status. The model also controls for a set of practice-level controls X'_p , in particular the specialty group, as well as a set of county-level controls Z'_c such as the county-level unemployment rate or the physician density (see Table 1). As mentioned, the model routinely controls for the day-of-the-week during which the caller called a practice (DOW_{ip}) as well as the time-of-the-day (TOD_{ip}) of the call.

In the saturated specifications, we add calendar-date fixed effects, δ_t . Similarly, we replace the practice-level controls with practice fixed effects ρ_p in some specifications. We routinely cluster the standard errors e_{ip} at the practice level and estimate linear probability models using OLS. (However, we also test the robustness of the coefficients using probit models and calculating marginal effects, Table A3.)

Our second model uses the balanced Sample B and is:

$$\ln(w_{ip}) = \alpha + \beta PHI_i + X'_p \tau + Z'_c \theta + \gamma DOW_{ip} + \delta_t + \rho_p + e_{ip} \quad (2)$$

where w_{ip} stands for our second outcome variable *dayswait*, and measures the wait time in weekdays for fictional patient i in practice p . It is continuous but skewed to the left (Figure A2), which is one reason why we replace 0s with 0.01 and take the logarithm. The coefficient estimates of the main variable of interest, PHI_i , then approximate the wait time differential between publicly and privately insured patients in percent. The other control variables are defined as above. We also test the robustness of the results by using the plain w_{ip} variable and running negative binomial count data models that consider excess zeros and overdispersion.

In extended specifications, we test for effect heterogeneity by interacting PHI_i with regional and other variables and add these interaction terms to the model.

7. Results

Nonparametric Findings

We start by plotting nonparametric results. In a perfectly randomized setting, they should very well approximate the parametric findings that additionally control for date, day-of-week, time-of-day and practice fixed effects.

Figure 2 plots bar diagrams of the first outcome variable *apptm* along with 95% confidence intervals. As can be seen with bare eyes, the share of privately insured who were offered an appointment (88%) is larger than the share of publicly insured who were offered an appointment (83%). A Fisher's exact test shows that this difference is significant at the 5% level (Fisher's exact = 0.002) level.

[Figure 2 about here]

Next, Figure 3 plots the distribution of the second outcome variable *dayswait* separately for the privately and publicly insured using the balanced Sample B. Again, it is easy to see that the wait time distribution for the privately insured is much more left-skewed than the wait time distribution for the publicly insured. The former has a lot more mass over the 0 to 20-weekday support region, and the latter has more mass exceeding 20-weekdays of wait time as well as a much longer right tail.

[Figures 3 and 4 about here]

Figure 4 plots bar diagrams along with 95% confidence intervals to illustrate mean differences in wait times between the publicly and privately insured, also using Sample B. As seen, the mean wait time for publicly insured is almost twice as long and 25 weekdays, whereas the mean wait time for

privately insured is below 12 weekdays. The confidence intervals clearly do not overlap, indicating a highly significant difference in wait times, depending on the insurance status. The difference also holds for the median wait time, which is 15 for publicly insured and 6 for privately insured patients. This prior is confirmed by a two-sided Wilcoxon Signed-Rank Test which is significant at the 0.1% level.²³

[Figures 5 and 6 about here]

Finally, Figures 5 and 6 show the same bar diagrams for *apptm* and *dayswait* but separately for the three specialist groups. In conjunction with the differences in reimbursement rates (Table 1), this heterogeneity test links cream-skimming to actual differences in patient profitability. On the extensive margin, we do not find much evidence for systematic patient cream skimming (although privately insured have always higher appointment rates which are, however, not statistically significant). By contrast, on the intensive margin, we do find evidence for systematic patient cream skimming and significant wait time differences for gastroscopies, allergy tests and hearing tests. Note that the differences are much smaller for hearing tests where reimbursement rates are much lower and the absolute and relative price differences much smaller (€16 vs. 23).

Parametric Findings

Next, we move on to our parametric findings and multivariate regression models. Table 3 shows the findings from our first model in equation (1), which uses the binary *apptm* measure as outcome variable. These models assess the effect of insurance status on being offered an appointment for non-

²³ Figure A3 in the Appendix shows the cumulative density functions (cdf) of wait time in weekdays for all successfully contacted practices that offered an appointment under both insurance types (i.e. Sample B). The cdf of privately insured patients dominates the cdf of publicly insured patients over the entire region of support. Note that the existence of practices that only offered an appointment for privately insured patients as shown in Figure 2 (33 practices) exacerbate the differences in waiting times between privately and publicly insured patients.

urgent treatments. Each column in Table 3 represents one model. The models only differ by the inclusion of different sets of covariates as indicated in the bottom panel of the Table.

[Table 3 about here]

Table 3 shows the following: First, we find that the insurance status of the caller has a highly significant impact on the willingness of practices to schedule an appointment. Two coefficient estimates are significant at the 1% level, and two are significant at the 5% level. Being privately insured increases the likelihood of an appointment by 3.8 to 5.2 percentage points or by between 4.2 and 5.7% relative to the mean. Second, the estimates are robust across all model specifications and not statistically different from one another. The inclusion of week-of-year fixed effects, county fixed effects, and even practice fixed effects barely alter the coefficients although the inclusion of specialty controls slightly increases the effect sizes and the practice fixed effects slightly decreases them. Third, the findings are also robust to running probit models and calculating marginal effects (Table A3).

Table 4 follows the same setup as Table 3 but estimates our second model and equation (2). It is basically identical to equation (1), but uses the second continuous outcome variables *dayswait*, which counts the wait times in weekdays. The coefficient estimates then indicate the effect of a private insurance status on the mean wait time in weekdays, relative to publicly insurance status.²⁶ In the Appendix, in Table A4, we replicate Table 4 with count data models and in Table A5, we do not take the logarithm of the dependent variable.²⁷ All eight models in Table 4, A4 and A5 use our balanced Sample

²⁶ Table 4 uses the logarithm of the wait time in weekdays as dependent variable. For example, in our preferred model in column (4), the difference in wait times in percent would be $\exp(-1.1381)-1=-0.6796$ or about -68%. In contrast, Figure 4 shows the unconditional mean differences which are 24.89 vs. 11.57 days wait time or a $(13.32/24.89)*100=54\%$ difference. See Section 6 for a discussion of the randomization process and why the regression models control for seasonal, regional and other covariates.

²⁷ Table A6 in the Appendix replicates the main regression considering only the first successful contact for each practice (Panel A) and considering only the three main specialist groups (Panel B). As seen in Panel A, restricting the sample to the

B and only include the 502 unique practices that offered specific appointments to both callers, the publicly and the privately insured.

[Table 4 about here]

Tables 4, A4 and A5 show the following: All 12 model coefficients are highly significant. The estimates are very robust to the sets of covariates included, reinforcing that our randomization was successful. This also implies the absence of structural differences in terms of the week-of-the-year or the county of residence. Moreover, because the differences are very close to the differences of simple nonparametric tests, it also suggests the absence of structural imbalances by day-of-the-week or the time of the day of the call.

The results show that privately insured patients wait on average 13 fewer weekdays for an appointment, conditional on being offered one. In other words, publicly insured patients have to wait more than twice as long for an appointment; the mean wait time for the publicly insured is 24.9 weekdays (or about 5 weeks on average), whereas the mean wait time for privately insured patients is only 11.6 weekdays (or a little more than 2 weeks on average).

Differences by Specialist Group and Profitability

We already briefly discussed differences by specialty above. Recall that we know the list prices for specific SHI and PHI treatments as well as the average claim amounts for ICD-10 code T78 (“allergies”) as well as K29 (“Gastritis and duodenitis”), K30 (“Functional dyspepsia”) or K31 (“Other diseases of stomach and duodenum”). It is clear from Table 1 that the requested endoscopy and related

first successful contact provides similar results than our baseline analysis but the coefficients for the extensive margin are slightly larger and the coefficients for the intensive margin are slightly. As seen in Panel B, considering only the three main specialties yields robust results as well but slightly smaller coefficients for both the extensive and intensive margin. Table A7 reports results for the unbalanced Sample A. The results are very similar compared to the baseline analysis.

diseases are significantly more profitable when treating privately insured. The same is true for allergy tests. Hearing tests, by contrast, show the smallest profitability differences both in relative and absolute terms. Table 5 shows our regression results for *apptm* and *dayswait* separately for the three specialist groups. Figures 5 and 6 show the analog nonparametric visual results in form of bar diagrams along with 95% confidence intervals.

[Table 5 about here]

As mentioned and seen in Panel A of Table 5 and Figure 5, when differentiating by specialty, although the points estimates are systematically positive and smaller for the less profitable hearing tests, they do not have enough statistical power to identify differences in a strict statistical sense.

However, as for the intensive margin in Panel B (*dayswait*) of Table 5 and Figure 6, the findings are very similar to the main analysis and statistically significant for each specialty. Most importantly, the effect size differences are largest for endoscopies and smallest for hearing tests, lining fully up with the reported differences in profitability in Table 1. However, although these results are highly suggestive, the differences are not different in a statistical sense due.

Effect Heterogeneity by County Characteristics

Next, we test for heterogeneity in inequality in health care access. Technically, we interact our variable of interest *Privately Insured* with one of the following stratifying county-level covariates: physician density, population density, household income, East Germany, and share of privately insured in state. Then we add the interaction term along with the two variables in levels to the models in

equations (1) and (2). Panel A of Table 6 shows the results for *apptm* and Panel B of Table 6 shows the results for *dayswait*.²⁸

[Table 6 about here]

As seen, only one of the interaction terms is statistically significant. The findings for physician density (column 1), household income (column 3), East Germany (column 4) and share of privately insured patients (column 5) are all far from being significant in both panels. The only significant finding in column (2) of Panel B suggests that a higher population density, e.g. in cities as compared to more rural counties, is associated with *more* discrimination and inequality in access.

Discussion

Our field experiment clearly shows that privately insured patients (i) are offered appointments at significantly higher rates and (ii) are offered appointments with shorter wait times, compared to publicly insured patients with lower reimbursement rates. This holds in the German context with its two-tier health care system where reimbursement rates structurally vary between the two systems but are otherwise identical across plans within each system. We also find strongly suggestive evidence that the discrimination is larger when the reimbursement rate differences are larger, e.g., for gastroscopies (upper gastrointestinal endoscopy), and smaller when the reimbursement rate differences are smaller, e.g., for hearing tests.

However, as we randomized the insurance status of the fictional patient in our audit study, our “cream-skimming” interpretation also allow for alternative mechanisms. However, we believe that all specific alternative explanations can ultimately be subsumed as: *practices structurally select more*

²⁸ As a robustness check, Table A8 shows the same heterogeneity analyses for only the first successful contacts.

profitable patients:²⁹First, doctors may not only expect a significantly higher reimbursement but also deal with healthier patients who have fewer co-morbidities (Table A1). Given a specific diagnosis, healthier patients imply shorter treatments and interactions and are thus more profitable patients. On the other hand, comorbidities and the health status are less relevant for highly standardized treatments like hearing tests, allergy tests or gastroscopies.

Second, one could hypothesize that doctors prefer better-educated patients because of the patient's education itself. Some studies have shown that doctors prefer patients with a good job or high socioeconomic status (Olah et al. 2013; Angerer et al. 2019). We would argue, however, that a high socio-economic status is simply a proxy for more profitable, wealthy patients (and not vice versa).³⁰

Third, in Germany, the privately insured have on average higher incomes (Table A1) and their reimbursement is 100% fee-for-service (Section 3). Hence, it may not just be the very narrowly defined one-time reimbursement rate differential in Table 1, but a broader definition of profitability that makes doctors cream-skim the privately insured. For example, doctors may schedule profitable follow-up visits or sell additional, potentially medically unnecessary and not covered services to the more affluent privately insured. On the other hand, SHI plans have basically zero cost-sharing and public insurers also directly pay providers without intensive claim reviews.³¹ Private insurers usually share costs, review

²⁹ Note that, in Germany, no official quotas or rules for how to provide appointments for publicly and privately patients exist. Essentially, absent emergency cases, private practices determine how they schedule appointments. Unlike in the U.S., e.g. for Medicaid patients, 90% of the population are publicly insured and not seen as charity cases. Rather, lots of anecdotal evidence suggests that the privately insured are regarded as particularly profitable and thus offered faster appointments and longer time slots during appointments (Soester Anzeiger 2019).

³⁰ Moreover, there is anecdotal evidence that doctors consider highly-educated, "empowered" patients as rather annoying, time-consuming and difficult (Neuberger 2000; Rankin 2011).

³¹ In Germany, the review of SHI-claims is mainly the physician associations' task (see § 106d of the social code book V (SGB V)). In practice, these reviews are probably less strict for the publicly insured than for private patients where PHI directly reviews claims and have a strong(er) incentive to deny reimbursement. E.g. the German Health Expert Council

claims, and do not pay providers directly; patients do and then submit the claims to their insurer. Hence, the risks of claim disputes and non-payments are higher when treating privately insured.

Finally, as shown in Figure 6, inequality in access is least pronounced for hearing tests where the absolute reimbursement rates are the lowest, and the relative difference between private and public rates is also the smallest (Table 1). To us, the findings by specialist groups and their implied specific treatment profitability strongly suggest that the specific reimbursement rate differentials do drive cream-skimming of the more profitable privately insured.³²

Although our research identifies drivers of structural inequalities in health care access, we deliberately abstain from drawing welfare conclusions. While structural differences in reimbursement rates could be efficient, well-informed consumers as well as consumer choice are two important ingredients for efficiency (cf. Handel and Kolstad 2015; Bhargava, Loewenstein and Sydnor 2017).³³ In the German case, 77% of the population do not have the option to opt out of the public insurance and purchase only private insurance, which provides faster and better access. In the U.S. case, while consumers theoretically have the “choice” between employer-sponsored private coverage, Medicaid and Medicare, all insurance schemes are inherently intertwined with equality in opportunity, race,

„Sachverständigenrat Gesundheit“ documented very low recourse claims for SHI-treatments: In 2008, only about 0.5 % of practices had to pay money back due submitting incorrect claims (SVR Gesundheit 2014).

³² While it is most likely that the identified access differences by specialty are driven by the reimbursement rate differences, there is the possibility that these are instead driven by selection into specialties and associated unobserved practice differences. For example, Li (2018) shows that medical students with lower degrees of altruism tend to select specialties with the highest expected income in the United States. In Germany, however, these income differences are much smaller.

³³ Another argument by German private insurers and doctor representatives is that the higher private rates help cross-subsidizing the lower public rates. Although it may be true in single cases that doctors would have to give up their practice without the privately insured, it is also true that the medical profession is the occupational group with the lowest unemployment rate and highest average incomes. While rural areas lack specialists and primary care physicians, policy reforms that substantially increased reimbursement rates in those areas did not lead to a strong increase in the supply of doctors in those regions (SVR Gesundheit 2014). Insolvencies of physicians are low in Germany (Destatis 2019); thus, private patients do not seem to be instrumental for the economic survival of private practices in Germany.

poverty and age. As mentioned in the introduction, inequality in access to health care is just *one* dimension on which to rate health care systems. Trade-offs with other dimensions can and do exist. However, in our opinion, this should not preclude health economists from studying this important dimension, just as studying equality in incomes and wealth is an independent topic of inquiry.

8. Conclusion

The main objective of this research was to implement a field experiment to assess the impact of public-private insurance status and related reimbursement rate differences on health care access. We use the German institutional setting for the field experiment because it is particularly well-suited as a clean testing ground. Germany is one of the very few countries with coexisting public and private insurance systems and structurally varying provider reimbursement rates between the two systems. The reimbursement rates for the privately insured are about two to three times higher and have classic fee-for-service schemes without caps or bundled payments. Importantly, reimbursement rates for both systems are standardized and centrally set, instead of being determined by negotiations between single insurers and providers. Provider networks and managed care do not exist in Germany and hence do not operate as a confounding factor in our setting.

Our test person called almost one thousand private specialist practices over the course of one calendar year. The test person followed a strict protocol and revealed the randomized insurance status of the fictional patient during the call, as common in Germany. The fictional patient called each practice twice; once pretending to be a publicly insured patient and once pretending to be a privately insured patient. In each case, the test person asked for an appointment for a non-urgent medical treatment: gastroscopies with gastroenterologists, hearing tests with otorhinolaryngologists, and allergy tests with allergists.

Our findings show that structural inequalities in reimbursement rates create structural inequalities in health care access. We document higher access barriers for less profitable patients, both on the extensive and intensive margin. Publicly insured patients were between four and six percent less likely to be offered an appointment. Moreover, when offered appointments, publicly insured patients had to wait 13 weekdays longer (more than twice as long) than privately insured patients. While one could argue that a three week longer wait time for specific population subgroups should not be reason for concern in non-urgent settings, recall that these patients may suffer three weeks longer due to undiagnosed allergies, hearing or stomach problems. Importantly, the main objective of this paper is to study driving forces of inequalities in health care access, not the health effects of such inequalities.³⁴ What's more, despite the fact that only a small percentage of the population is privately insured, most Germans seem to find the structural differences between the two systems unacceptable. In a representative survey among Germans, two thirds indicate that they would be concerned about a public-private "two-class" health care system (DHBW Mosbach 2014); indeed, the equalization of reimbursement rate differences is on the political agenda in Germany (Wasem und Walendzik 2019, Handelsblatt 2018).

The policy implications of our findings suggest that uniform reimbursement rates would reduce inequalities in health care access. Because, in the U.S. and Germany, healthier and wealthier individuals tend to have private insurance with higher reimbursements rates, such a system exacerbates structural inequalities in health care access and population health. However, uniform reimbursement rates may have unintended consequences and could result in welfare losses if they reduce the overall supply of

³⁴ At least for stomach problems, there exists hard evidence that health care access or long waiting times have adverse health effects. E.g. a treatment delay of more than one month for gastric cancer—which might be diagnosed by gastroscopy—is associated with higher mortality rates (Yun et al. 2012). If a person receives treatment before the stomach cancer spreads, the 5-year survival rate is 68 %; once it reaches distant organs, survival rates drop to 5%.

physicians or increase taxes or premiums. How optimal reimbursement rates should be set in a health care system in practice is an important topic for future research.

References

- Alexander, D., Schnell, M. 2017. Closing the gap: The impact of the Medicaid primary care rate increase on access and health, FRB of Chicago Working Paper No. WP-2017-10.
- Angerer, S., Waibel, C. Stummer, H. 2019. Discrimination in health care: A field experiment on the impact of patients' socio-economic status on access to care, *American Journal of Health Economics*, 5(4): 407-427.
- Asplin, B. R., Rhodes, K. V., Levy, H., Lurie, N., Crain, A. L., Carlin, B. P., Kellermann, A. L. 2005. Insurance status and access to urgent ambulatory care follow-up appointments, *Journal of the American Medical Association*, 294(10): 1248-54.
- Atal, J., H. Fang, M. Karlsson, Ziebarth, N. R. 2019. Exit, voice or loyalty? An investigation into mandated portability of front-loaded private health plans, *Journal of Risk and Insurance*, 86(3): 697-727.
- Baker, L. C., Royalty, A. B. 2000. Medicaid policy, physician behavior, and health care for the low-income population, *Journal of Human Resources*, 35(3): 480-502.
- Bauhoff, S. 2012. Do health plans risk-select? An audit study on Germany's social health insurance, *Journal of Public Economics*, 96(9-10): 750-759.
- BBSR - Bundesinstitut für Bau-, Stadt- und Raumforschung 2018. *Inkar. Indikatoren und Karten zur Raum- und Stadtentwicklung*. <http://www.inkar.de> (date of access: June 10, 2018).
- Bhargava, S., Loewenstein, G., Sydnor, J. 2017. Choose to lose: Health plan choices from a menu with dominated option, *The Quarterly Journal of Economics*, 132(3): 1319-1372.
- BMG – Bundesministerium für Gesundheit 2018. *KM 6-Statistik (gesetzliche Krankenversicherung: Versicherte)*, in: www.gbe-bund.de (date of access: July 8, 2019)
- BMG – Bundesministerium für Gesundheit 2019. *Statistik über Versicherte, gegliedert nach Status, Alter, Wohnort und Kassenart 2018*, in: <https://www.bundesgesundheitsministerium.de/themen/krankenversicherung/zahlen-und-fakten-zur-krankenversicherung/mitglieder-und-versicherte.html> (date of access: July 13, 2019)
- Brosig-Koch, J., Kairies-Schwarz, N., Kokot, J. 2017. Sorting into payment schemes and medical treatment: a laboratory experiment, *Health Economics*, 26(S3): 52-65.
- Bünnings, C., H. Schmitz, H. Tauchmann, Ziebarth, N.R. 2018. How health plan enrollees value prices relative to supplemental benefits and service quality, *Journal of Risk and Insurance*, 86(2): 415-449.
- Candon M., Zuckerman S., Wissoker D., Saloner, B. 2018. Declining Medicaid fees and primary care appointment availability for new Medicaid patients, *JAMA Internal Medicine*, 178(1):145-146.
- Chen, Y., Persson, P., Polyakova, M. 2019. The roots of health inequality and the value of intra-family expertise, NBER Working Papers 25618.
- CMS – Centers for Medicare and Medicaid Services 2018a. *Physician Fee Schedule*, in: <https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeeSched/index.html>, (date of access: September 10, 2019).
- CMS – Centers for Medicare and Medicaid Services 2018b. *CY 2018 Medicare Fee Schedule Changes*, in:

- <https://gi.org/wp-content/uploads/2018/01/MPFS-Final-2018-v-2017-Rates.pdf>, (date of access: September 10, 2019).
- Clemens, J., Gottlieb, J.D. 2014. Do physician's financial incentives affect medical treatment and patient health? *American Economic Review*, 104(4): 1320-1349.
- Commonwealth Fund, 2014. *Mirror, Mirror on the Wall 2014 Update: How the U.S. Health Care System Compares Internationally*, by Davis, K., Stremikis, K. Squires, D. and C.Schoen, in: <https://www.commonwealthfund.org/publications/fund-reports/2014/jun/mirror-mirror-wall-2014-update-how-us-health-care-system> (date of access: October 18, 2018).
- Commonwealth Fund, 2017. *Mirror, Mirror 2017. International Comparison Reflects Flaws and Opportunities for Better U.S. Health Care*, by Eric C. Schneider, Dana O. Sarnak, David Squires, Arnav Shah, and Michelle M. Doty, in: <https://www.commonwealthfund.org/publications/fund-reports/2017/jul/mirror-mirror-2017-international-comparison-reflects-flaws-and> (date of access: October 18, 2018).
- Cooper, C., Craig, S.V., Gaynor, M., Reenen, J.V. 2018. The price ain't right? Hospital prices and health spending on the privately insured, *The Quarterly Journal of Economics*, 134(1): 51–107.
- Cullis, J. G., Jones, P. R., Propper, C. 2000. Waiting lists and medical treatment: Analysis and policies, *Handbook of Health Economics*, 1(PART B), Chapter 23:1201-1249.
- Das Telefonbuch 2018. <https://telefonbuch.t-online.de/> (date of access: June 10, 2018).
- Decker, S. L. 2009. Changes in Medicaid physician fees and patterns of ambulatory care, *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, 46(3): 291-304.
- Decker, S.L., 2012. In 2011 Nearly one-third of physicians said they would not accept new Medicaid patients, but rising fees may help, *Health Affairs*, 31: 1673–1679.
- Destatis - Federal Statistical Office Germany 2019. *Unternehmen und Arbeitsstätteninsolvenzverfahren*, Fachserie 2 Reihe 4.1.
- Destatis - Federal Statistical Office Germany 2018a. *Daten aus dem Gemeindeverzeichnis - Kreisfreie Städte und Landkreise nach Fläche und Bevölkerung auf Grundlage des Zensus 2011 und Bevölkerungsdichte*. <https://www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis/Administrativ/05-staedte.html> (date of access: July 12, 2019).
- Destatis - Federal Statistical Office Germany 2018b. *Daten aus dem Gemeindeverzeichnis - Bundesländer mit Hauptstädten nach Fläche und Bevölkerung auf Grundlage des Zensus 2011 und Bevölkerungsdichte*, in: <https://www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis/Administrativ/02-bundeslaender.html> (date of access: July 12, 2019)
- Destatis - Federal Statistical Office Germany 2018c. *Unternehmen und Arbeitsstätten: Kostenstruktur bei Arzt- und Zahnarztpraxen sowie Praxen von psychologischen Psychotherapeuten*, Fachserie 2 Reihe 1.6.1.
- DHBW Mosbach 2014. *GfK-Umfrage: Zwei-Klassen-Medizin in Deutschland?*, Deutsche Hochschule Baden-Württemberg Mosbach, in: <https://www.mosbach.dhbw.de/news-1/news-im-detail/gfk-umfrage-zwei-klassen-medizin-in-deutschland?> (date of access: July 12, 2019)
- Ellis, R. P., McGuire, T. G. 1986. Provider behavior under prospective reimbursement, *Journal of Health Economics*, 5(2): 129-151.
- Felder, S. 2008. To wait or to pay for medical treatment? Restraining ex-post moral hazard in health insurance, *Journal of Health Economics*, 27(6): 1418-1422.
- GKV Spitzenverband 2018. *Statutory Health Insurance*, in: https://www.gkv-spitzenverband.de/english/statutory_health_insurance/statutory_health_insurance.jsp, (date of access: September 10, 2018).

- Goebel, J., Grabka, M. M., Liebig, S., Kroh, M., Richter, D., Schröder, C., Schupp, J. 2019. The German Socio-Economic Panel Study (SOEP), *Jahrbücher für Nationalökonomie und Statistik / Journal of Economics and Statistics*, 239(2): 345-360.
- GOÄ - Gebührenordnung für Ärzte 2019. <https://diego.one/go%C3%A4> (date of access: September 10, 2019).
- Gravelle, H., Siciliani, L. 2008. Optimal quality, waits and charges in health insurance, *Journal of Health Economics*, 27(3): 663-674.
- Halpern, M. T., Romaire, M. A., Haber, S. G., Tangka, F. K., Sabatino, S. A., Howard, D. H. 2014. Impact of state-specific Medicaid reimbursement and eligibility policies on receipt of cancer screening, *Cancer*, 120(19): 3016-3024.
- Handel, B. R., Kolstad, J. T. 2015. Health insurance for "humans": Information frictions, plan choice, and consumer welfare, *American Economic Review*, 105(8): 2449-2500.
- Handelsblatt 2018. Krankenkassen wettern gegen einheitliche Arzthonorare, in: <https://www.handelsblatt.com/politik/deutschland/gkv-und-pkv-krankenkassen-wettern-gegen-einheitliche-arzthonorare/20905326.html> (date of access: September 13, 2019)
- Heinrich, N., Wübker, A., Wuckel, C. 2018. Waiting times for outpatient treatment in Germany: New experimental evidence from primary data, *Jahrbücher für Nationalökonomie und Statistik / Journal of Economics and Statistics*, 238(5): 375-394.
- Holgash, K., Heberlein, M., 2019. Physician Acceptance Of New Medicaid Patients: What Matters And What Doesn't [WWW Document]. *Health Affairs Blog*, in: <https://www.healthaffairs.org/doi/10.1377/hblog20190401.678690/full/> (date of access: July 28, 2020)
- Hullegie, P. and Klein, T.J. (2012). The effect of private health insurance on medical care utilization and self-assessed health in Germany, *Health Economics*, 19(9): 1048-1062.
- Jones, A.M., Rice, N., Robone, S., Dias, P.R. 2011. Inequality and polarisation in health systems' responsiveness: A cross-country analysis, *Journal of Health Economics*, 30(4): 616-625.
- Jürges, H. 2009. Health Insurance Status and Physician Behavior in Germany, *Schmollers Jahrbuch : Journal of Applied Social Science Studies*, 129(2): 297-307.
- Karlsson, M., Klein, T., Ziebarth, N. R. 2016. Skewed, persistent and high before death: Medical spending in Germany, *Fiscal Studies (Special Issue: Medical Expenditures around the World)*, 37(3-4): 527-559
- Kesternich, I., Schumacher, H., Winter, J. (2015). Professional norms and physician behavior: Homo oeconomicus or homo hippocraticus? *Journal of Public Economics*, 131, 1-11.
- KBV - Kassenärztliche Bundesvereinigung 2017. *Einheitlicher Bewertungsmaßstab (EBM)*, in: http://www.kbv.de/media/sp/EBM_Gesamt_Stand_1_Quartal_2017.pdf (date of access: June 10, 2018).
- KBV - Kassenärztliche Bundesvereinigung 2018a. *Aktuelles zu den Honorarverhandlungen 2019*, in: <http://www.kbv.de/html/2054.php> (date of access: September 10, 2018).
- KBV - Kassenärztliche Bundesvereinigung 2018b. *Einheitlicher Bewertungsmaßstab (EBM)*, in: http://www.kbv.de/media/sp/EBM_Gesamt_Stand_2_Quartal_2018.pdf (date of access: December 09, 2019).
- Kuchinke, B., Sauerland, D., Wübker, A. 2009. The influence of insurance status on waiting times in German acute care hospitals: An empirical analysis of new data, *International Journal for Equity in Health*, 8(44).
- KVNO - Kassenärztliche Vereinigung Nordrhein, 2018. *Honorarverteilungsmaßstab der Kassenärztlichen Vereinigung Nordrhein*, in: <https://www.kvno.de/downloads/honorar/hvm1801.pdf> (date of access: June

10, 2018).

- Laudicella, M., Siciliani, M., Cookson, R. 2012. Waiting times and socioeconomic status: evidence from England, *Social Science & Medicine*, 74(9): 1331-1341.
- Levitt, S. D., List, J. A. 2007. What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives*, 21(2): 153-174.
- Levitt, S. D., List, J. A. 2009. Field experiments in economics: The past, the present, and the future, *European Economic Review*, 53(1): 1-18.
- Li, J. 2018. Plastic surgery or primary care? Altruistic preferences and expected specialty choice of U.S. medical students¹, *Journal of Health Economics*, 62: 45-59.
- Lindsay, C.M., Feigenbaum, B. 1984. Rationing by waiting lists, *The American Economic Review*, 74(3): 404-417.
- Lüngen, M., Stollenwerk, B., Messner, P., Lauterbach, K. W., Gerber, A. 2008. Waiting times for elective treatments according to insurance type: A randomized empirical study in Germany, *International Journal for Equity in Health*, 7.
- Ramos, A.L., Hoffmann, F. and Spreckelsen, O. 2018. Waiting times in primary care depending on insurance scheme in Germany, *BMC Health Services Research*, 18(1): 191-199.
- McGuire, T. G. 2000. Physician agency. In: *Handbook of Health Economics*, Eds. Culyer, A. J., Newhouse J. P., 1(Part A) Chapter 9: 461-536. Amsterdam: North Holland.
- Mackillop, William J., Zhou, Y., Quirt, C. F. 1995. A comparison of delays in the treatment of cancer with radiation in Canada and the United States, *International Journal of Radiation Oncology*, 32(2): 531-539.
- Maclean, J. C., McClellan C., Pesko, M. F., Polsky, D. 2018. Reimbursement rates for primary care services: Evidence of spillover effects to behavioral health, NBER Working Papers 24805.
- Monstad, K., Birger, L., Åeter, E., Espehaug, B. 2014. Waiting time and socioeconomic status: an individual-level analysis, *Health Economics*, 23(4): 446-461.
- Neuberger, J. 2000. The educated patient: new challenges for the medical profession, *Journal of Internal Medicine*, 247: 6-10.
- Nicholson S., Pauly M.V., Wu A.Y.J., Murray J.F., Teutsch S.M., Berger M.L. 2018. Getting real performance out of pay-for-performance, *The Milbank Quarterly*, 86(3): 435-457.
- Olah, M.E., Gaisano, G., Hwang, S.W. 2013. The effect of socioeconomic status on access to primary care: An audit study, *CMAJ*, 185(6): E263-E269.
- Pilny, A., Wübker, A., Ziebarth, N. R. 2017. Introducing risk adjustment and free health plan choice in employer-based health insurance: Evidence from Germany, *Journal of Health Economics*, 56(C): 330-351.
- Polsky, D., Richards, M., Basseyn, S., Wissoker, D., Kenney, G M., Zuckerman, S., Rhodes, K. V. 2005. Appointment availability after increases in Medicaid payments for primary care, *New England Journal of Medicine*, 372(6): 537-545.
- Rankin, L. 2011. When Docs Get Annoyed At Empowered Patients, *Psychology Today*, Jul 19, 2011, in: <https://www.psychologytoday.com/us/blog/owning-pink/201107/when-docs-get-annoyed-empowered-patients>, (date of access: July 7, 2019).
- Reif, S., Wichert, S., Wuppermann, A. (2018): Is it good to be too light? Birth weight thresholds in hospital reimbursement systems, *Journal of Health Economics*, 59: 1-25.
- Reinhardt, U. E. 2011. The many different prices paid to providers and the flawed theory of cost shifting: is it time for a more rational all-payer system? *Health Affairs*, 30(11): 2125-2133.

- Resneck, J., Pletcher, M.J., Lozano, N., 2004. Medicare, Medicaid, and access to dermatologists: The effect of patient insurance on appointment access and wait times, *Journal of the American Academy of Dermatology*, 50: 85–92.
- Riach, P. A., Rich, J. 2004. Deceptive field experiments of discrimination: Are they ethical? *Kyklos*, 57(3): 457-470.
- Richards, M. R., Polsky, D. 2016. Influence of provider mix and regulation on primary care services supplied to U.S. patients, *Health Economics, Policy and Law*, 11(2): 193-213.
- Roll, K., Stargardt, T., Schreyögg, J. 2012. Effect of type of insurance and income on waiting time for outpatient care, *The Geneva Papers on Risk and Insurance Issues and Practice*, 37(S4): 609-632.
- Sachverständigenrat zur Begutachtung der Entwicklung im Gesundheitswesen 2014. *Bedarfsgerechte Versorgung - Perspektiven für ländliche Regionen und ausgewählte Leistungsbereiche*, Gutachten 2014 Taschenbuch – 1. Oktober 2014.
- Saloner, B., Polsky, D., Friedman, A., Rhodes, K. 2015. Primary care appointment availability and preventive care utilization: Evidence from an audit study, *Medical Care Research and Review*, 72(2): 149-167.
- Schmitz, H. (2013). Practice budgets and the patient mix of physicians – The effect of a remuneration system reform on health care utilisation, *Journal of Health Economics*, 32(6): 1240-1249.
- Schmitz, H., Ziebarth, N. R. 2017. Does framing prices affect the consumer price sensitivity of health plan choice? *The Journal of Human Resources*, 52(1): 89–128.
- Schwierz, C., Wübker, A., Wübker, A., 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.
- Simon, M., 2017. *Das Gesundheitssystem in Deutschland: Eine Einführung in Struktur und Funktionsweise*. Bern: Hogrefe Verlag.
- Siciliani, L., Hurst, J. 2005. Tackling excessive waiting times for elective surgery: A comparative analysis of policies in 12 OECD countries, *Health Policy*, 72(2): 201-215.
- Siciliani, L. 2006. A dynamic model of supply of elective surgery in the presence of waiting times and waiting lists, *Journal of Health Economics*, 25(5): 891-907.
- Skaggs, D.L., Clemens, S.M., Vitale, M.G., Femino, J.D., Kay, R.M., 2001. Access to Orthopedic Care for Children With Medicaid Versus Private Insurance in California, *Pediatrics*, 107: 1405–1408.
- Sommers, B. D., McMurtry, C. L., Blendon, R. J., Benson, J. M., Sayde, J. M. 2017. Beyond health insurance: remaining disparities in U.S. health care in the post-ACA era, *The Milbank Quarterly*, 95(1): 43-69.
- Socio-Economic Panel (SOEP): Data for years 1984-2016, version 33, SOEP, 2019, doi: 10.5684/soep.v33.
- Soester Anzeiger 2019. Warum Patienten lange auf einen Facharzt-Termin warten, in: <https://www.soester-anzeiger.de/lokales/soest/manche-fachaerzte-lehnen-neue-patienten-12927558.html>, (date of access: March 13, 2020).
- Spiegel 2015. Lügen für schnellen Arzttermin "Klar bin ich privat versichert", in: <https://www.spiegel.de/gesundheit/diagnose/arzttermin-darf-man-sich-als-privatpatient-ausgeben-a-1064882.html> (date of access: July 28, 2020).
- Sundmacher, L., Kopetsch, T. 2013. Waiting times in the ambulatory sector--the case of chronically ill patients, *International Journal for Equity in Health*, 12(1): 77-87.
- van Doorslaer, E., Wagstaff, A., van der Burg, H., Christiansen, T., De Graeve, D., Duchesne, I., Gerdtham, U.-G., Gerfin, M., Geurts, J., Gross, L., Häkkinen, U., John, J., Klavus, J., Leu, R. E., Nolan, B., O'Donnell, O., Propper,

C., Puffer, F., Schellhorn, M., Sundberg, G., Winkelhake, O. 2000. Equity in the delivery of health care in Europe and the US, *Journal of Health Economics*, 19(5): 553-583.

Versorgungsatlas 2018. *Versorgungsstrukturen*, in:

<https://www.versorgungsatlas.de/themen/versorgungsstrukturen/> (date of access: June 10, 2018).

Viberg, N., Forsberg, B. C., Borowitz, M., Molin, R. 2013. International comparisons of waiting times in health care - Limitations and prospects, *Health Policy*, 112(1-2): 53-61.

Walendzik, A. und Wasem, J. (2019), *Ärztliche Vergütung ambulanter und ambulant erbringbarer Leistungen Gesundheitspolitisch zielgerechte Integrationsmodelle über sektorale Leistungsträger und Finanzierungssysteme*, Bertelsmann Stiftung (Eds.). First edition, in: <https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/aerztliche-verguetung-ambulanter-und-ambulant-erbringbarer-leistungen/> (date of access: March 13, 2020).

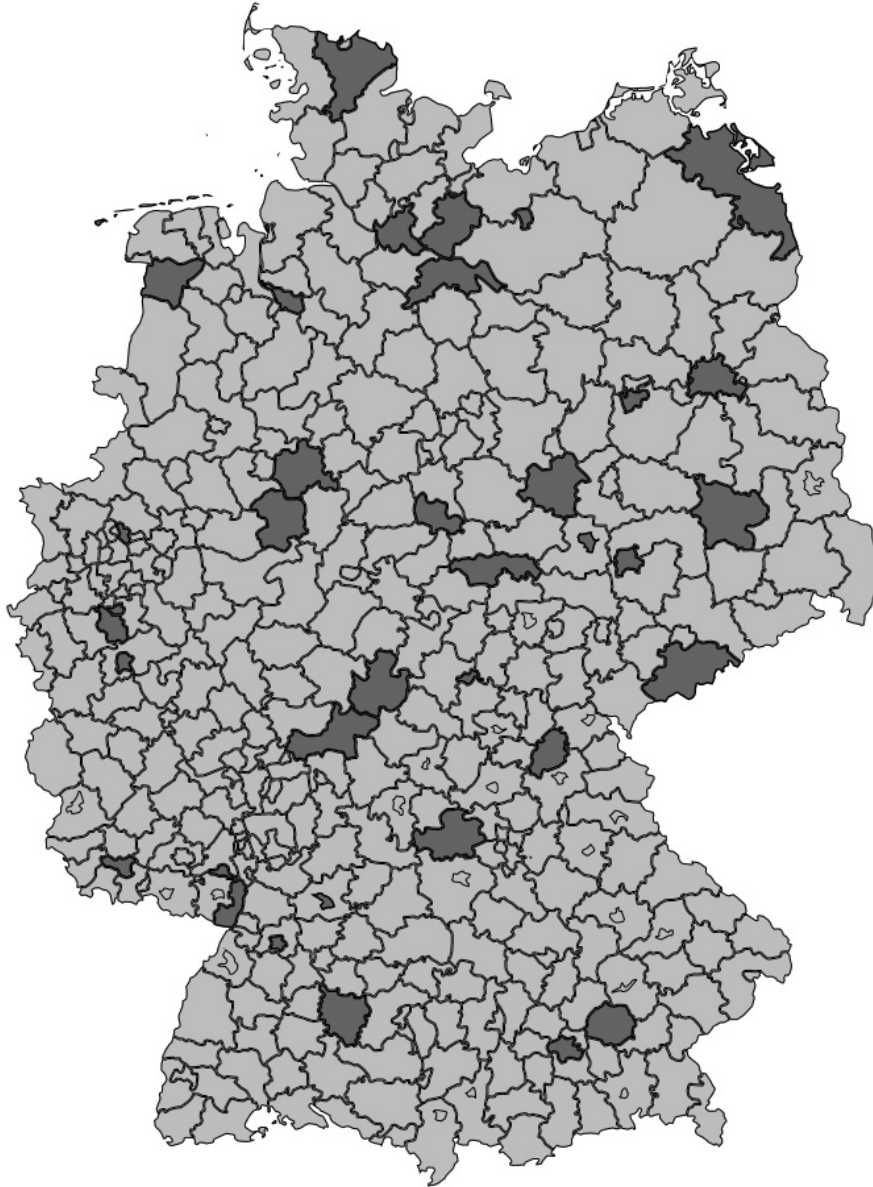
Walendzik, A., Greß, S., Manouguian, M., Wasem, J. 2008. Vergütungsunterschiede im ärztlichen Bereich zwischen PKV und GKV auf Basis des standardisierten Leistungsniveaus der GKV und Modelle der Vergütungsangleichung, *Diskussionbeitrag Universität Duisburg-Essen*, 165.

WIdO 2020. AOK-Abrechnungsdaten nach § 295 SGB V.

Yun, Y. H., Kim, Y. A., Min, Y. H., Park, S., Won, Y. J., Kim, D. Y., ... , Lee, D. H. 2012. The influence of hospital volume and surgical treatment delay on long-term survival after cancer surgery. *Annals of Oncology*, 23(10): 2731-2737.

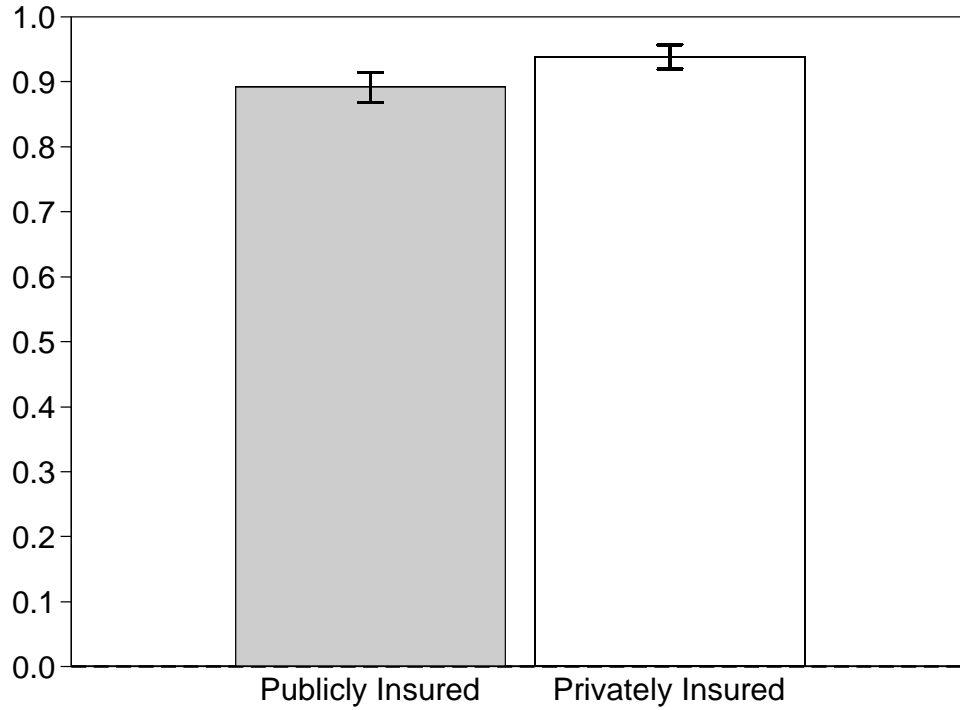
Figures and Tables

Figure 1: Selected Counties for Field Experiment



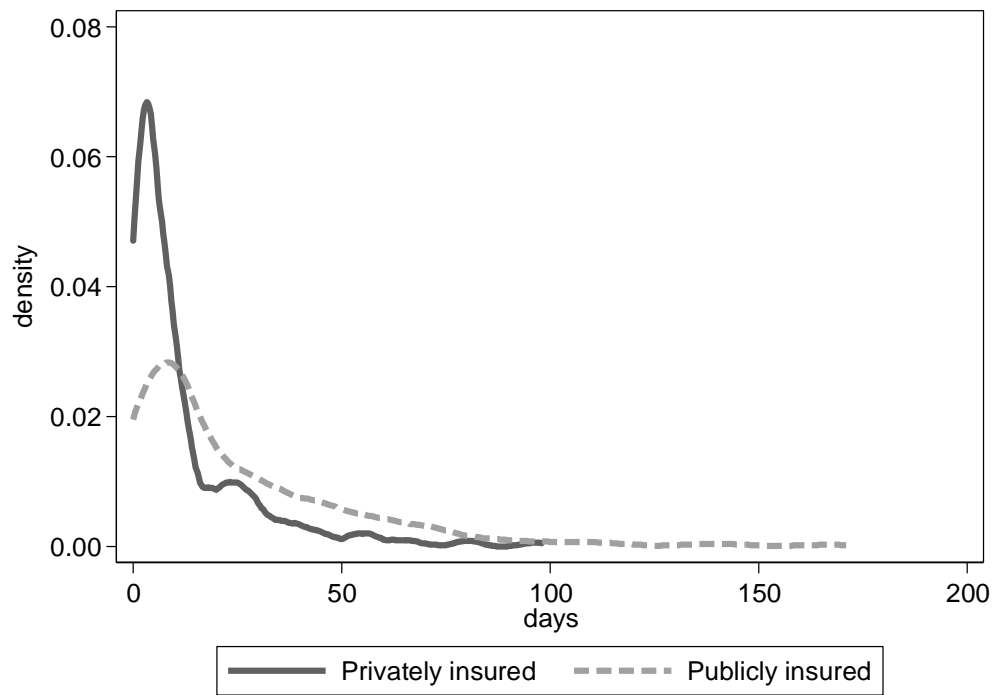
Source: Own illustration. The 36 counties selected for the field experiment are dark gray. Between one and four counties in each of the 16 federal states were selected (see Appendix B).

Figure 2: Likelihood to be Offered Appointment by Insurance Status



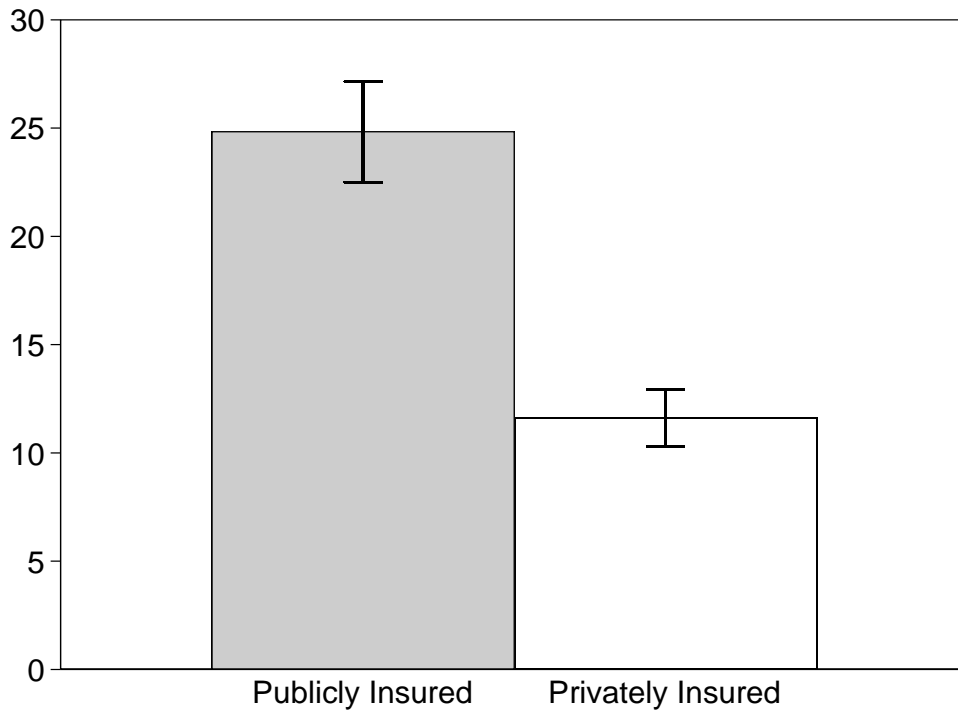
Source: Graph uses Sample A. The bars show 95% confidence intervals, derived from a regression of *apptm* on *Privately Insured* with standard errors clustered at the practice level.

Figure 3: Distribution of Wait Times in Weekdays by Insurance Status



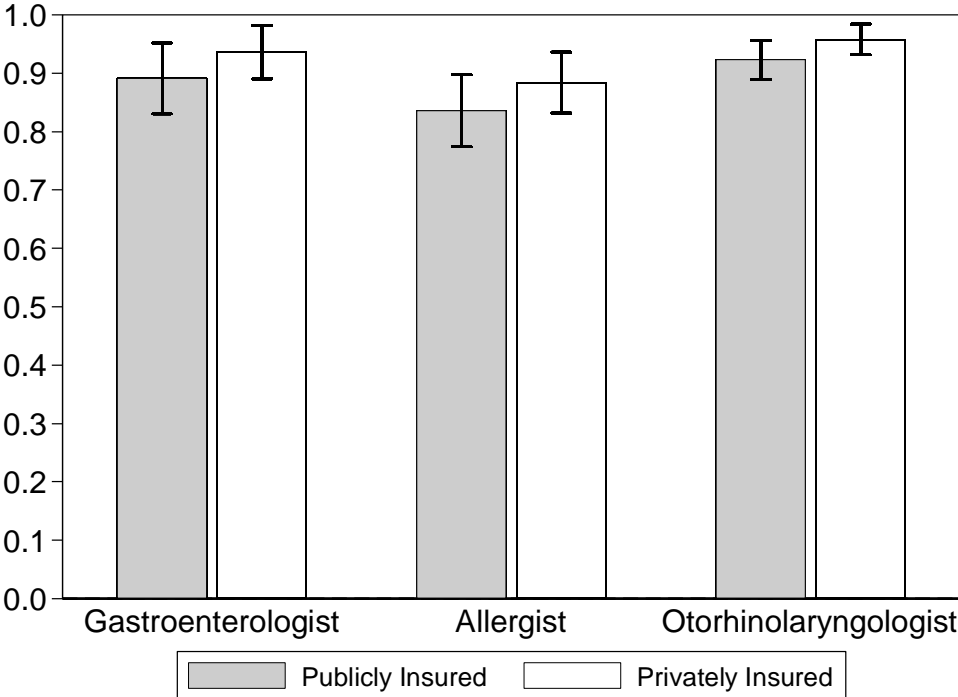
Source: Graph uses Sample B. X-axis shows the number of weekdays, counting from the day-of-the-call until an appointment was offered.

Figure 4: Average Wait Times by Insurance Status



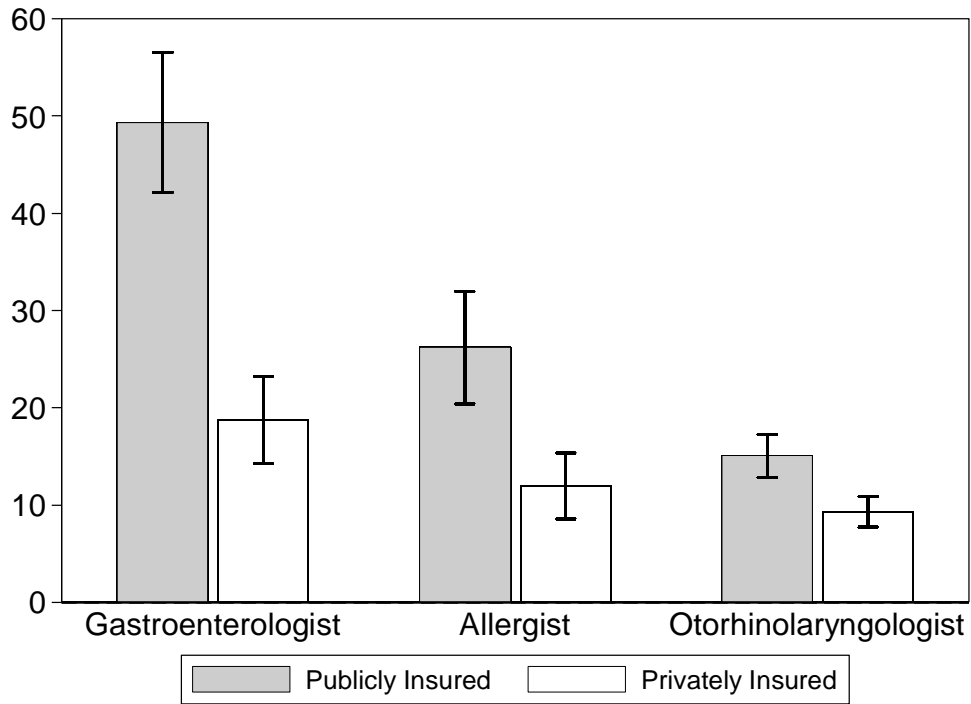
Source: Graph uses Sample B. The bars show 95% confidence intervals, derived from a regression of *dayswait* on *Privately Insured* with standard errors clustered at the practice level.

Figure 5: Likelihood to be Offered Appointment by Insurance Status and Specialist



Source: Graph uses Sample A. The bars show 95% confidence intervals, derived from separate regressions by specialty of *apptm* on *Privately Insured* with standard errors clustered at the practice level.

Figure 6: Average Wait Times by Insurance Status and Specialist



Source: Graph uses Sample B. The bars show 95% confidence intervals, derived from separate regressions by specialty of *dayswait* on *Privately Insured* with standard errors clustered at the practice level.

Table 1: SHI and PHI Reimbursement for Requested Treatments

	EBM			GOÄ		
	Points	Euros	Average claim amount per visit	Points	Euros	Average claim amount per visit
Allergy test	458	48.80	109.6	45	183.60	201.49
Hearing test	147	15.66	N/A	158	23.02	N/A
Gastroenterology	835	88.96	171.8	800	204.02	244.47

Notes: The first two columns display public reimbursement rates according to the *Einheitlicher Bewertungsmaßstab* (EBM) or “Unified Assessment Scale.” The next column uses claims data from the biggest public health insurer group in Germany (“Allgemeine Ortskrankenkassen”) with a market share of roughly 35 percent. It indicates the total claim amounts when the diagnosis contained ICD-10 code T78 (“allergies”) or ICD-10 codes K29 (“Gastritis and duodenitis”), K30 (“Functional dyspepsia”) or K31 (“Other diseases of stomach and duodenum”) (WIdO (2020)). The next two columns display private reimbursement rates according to the *Gebührenordnung für Ärzte* (GOÄ) or “Fee Schedule for Physicians” assuming an adjustment factor of 3.5. Values for GOÄ allergy tests are for 20 skin prick tests. The EBM does not differentiate between the number of prick tests. We used following GOÄ numbers: GOÄ 676 for gastroscopy, GOÄ 385 for allergy tests and GOÄ 1403 for hearing tests. GOÄ-code 676 is “Stomach examination under visual control using a camera to be used endogastrally, including photographs” and GOÄ-code 683 “Gastroscopy including oesophagoscopy using fully flexible optical instruments, including sample excision and/or puncture” (both gastroscopy). We used following EBM numbers: EBM 13400 for gastroscopy, EBM 30111 for allergy tests and EBM 20320 for hearing tests. The final column uses claims data from a big private insurer, which is roughly representative of the privately insured population in Germany (see Karlsson et al. 2016 for details). It indicates the total claim amounts when the diagnosis contained ICD-10 code T78 (“allergies”) or ICD-10 codes K29 (“Gastritis and duodenitis”), K30 (“Functional dyspepsia”) or K31 (“Other diseases of stomach and duodenum”); standard errors for the last column and “allergy test” is 450.24 and for the last column and Gastroenterology” is 372.38.

Table 2: Descriptive Statistics

VARIABLES	(1) mean	(2) sd	(3) min	(4) max	(5) N
<i>Dependent Variables</i>					
Apptm	0.915	0.279	0	1	1,328
Dayswait	19.01	23.17	0	171	1,215
Dayswait II	18.36	22.71	0	171	1,215
<i>Main Independent Variables</i>					
Privately Insured	0.501	0.500	0	1	1,328
Allergy test	0.215	0.411	0	1	1,328
Hearing test	0.363	0.481	0	1	1,328
Gastroscopy	0.159	0.366	0	1	1,328
First call privately insured	0.486	0.500	0	1	1,328
<i>County-Level Controls</i>					
Unemployment rate in %	7.924	2.634	2.100	13.60	1,328
Share Employees w/ academic degree in %	10.31	3.928	3.800	16.80	1,328
HH income per capita, €	1,789	318.2	1,362	3,451	1,328
Physician density	2.247	0.577	1.172	3.300	1,328
Residents per km ² in 1000	2,174	1,590	55	4,668	1,328
Eastern federal state	0.133	0.339	0	1	1,328
Share PHI ¹ in %	12.03	4.868	4.354	17.87	1,328

Sources: See Section 4 and 5 for details. County-level controls are taken from BBSR (2018) and Destatis (2018a, b).¹Available at state level.

Table 3: Impact of Insurance Status on Likelihood to be Offered Appointment

	(1) <i>apptm</i>	(2) <i>apptm</i>	(3) <i>apptm</i>	(4) <i>apptm</i>
Privately Insured	0.0483***	0.0515***	0.0511***	0.0384**
	(0.0117)	(0.0117)	(0.0119)	(0.0164)
Day-of-week FE	X	X	X	X
Week-of-year FE	X	X	X	X
Time of day	X	X	X	X
Specialty controls		X	X	X
Practice FE				X
County FE			X	
Observations	1,328	1,328	1,328	1,328
R-squared	0.0225	0.0585	0.1125	0.8506

Notes: Robust standard errors in parentheses, clustered at the practice level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Each column is one model as in equation (1) using Sample A, see Section 5 for details. The mean of the dependent binary variable *apptm* indicating whether an appointment was offered is 0.85 (see Table 1).

Table 4: Impact of Insurance Status on Wait Times

	(1) <i>Log(dayswait)</i>	(2) <i>Log(dayswait)</i>	(3) <i>Log(dayswait)</i>	(4) <i>Log(dayswait)</i>
Privately Insured	-1.0700*** (0.0871)	-1.1073*** (0.0885)	-1.1218*** (0.0891)	-1.1381*** (0.1276)
Day-of-week FE	X	X	X	X
Week-of-year FE	X	X	X	X
Time of day	X	X	X	X
Specialty controls		X	X	X
Practice FE				X
County FE			X	
Observations	1,004	1,004	1,004	1,004
R-squared	0.1665	0.2435	0.3058	0.7270

Notes: Robust standard errors in parentheses, clustered at the practice level. *** p<0.01, ** p<0.05, * p<0.1. Each column is one model as in equation (2) using Sample B, see Section 5 for details. The mean of the dependent variable *dayswait* indicating the number of weekdays until the offered appointment is 24.89 for publicly and 11.57 days privately insured patients (see Figure 4). The overall mean is 18.23 (all values for Sample B, not shown in Table 1). All models use the logarithm of *dayswait* where values of 0 have been replaced with 0.01.

Table 5: Impact of Insurance Status on Likelihood to be Offered Appointment and Wait Times

	(1) Gastroscopy	(2) Allergy test	(3) Hearing test
Panel A: <i>apptm</i>			
Privately Insured	0.0706 (0.0600)	0.0223 (0.0670)	0.0345 (0.0231)
Observations	211	286	482
R-squared	0.891	0.923	0.814
Panel B: <i>Log(dayswait)</i>			
Privately Insured	-1.482*** (0.341)	-1.225*** (0.314)	-0.834*** (0.206)
Observations	142	180	402
R-squared	0.790	0.779	0.725

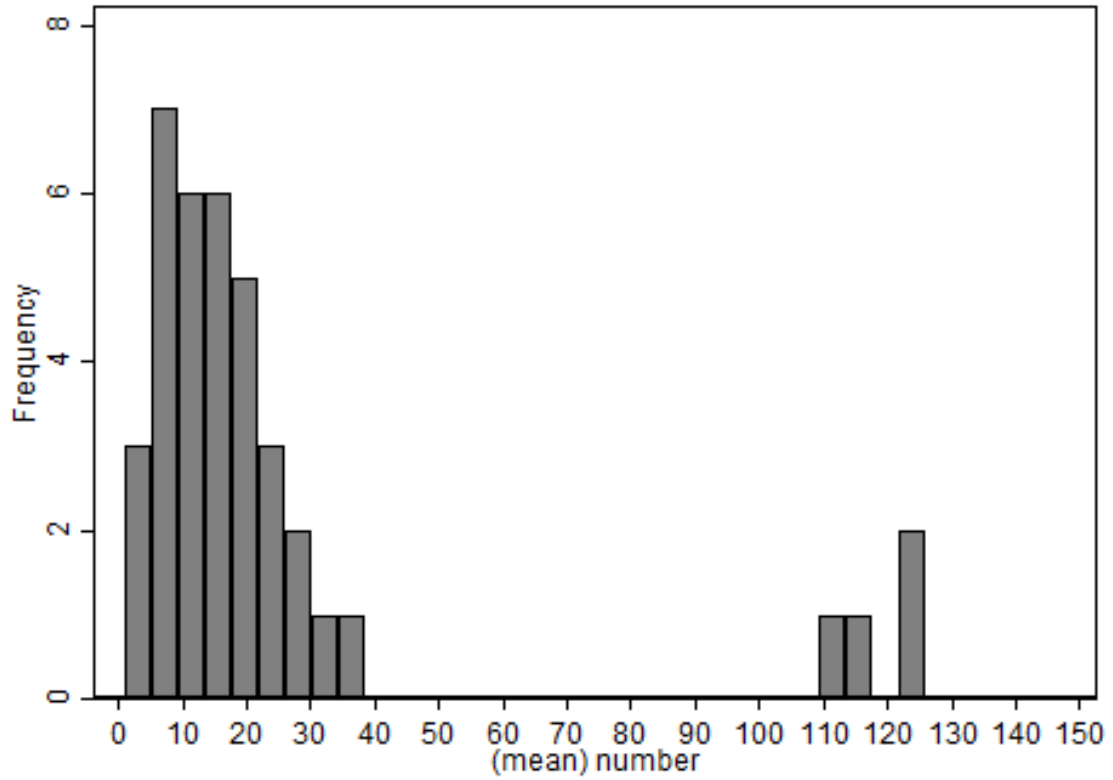
Notes: Robust standard errors in parentheses, clustered on practice level *** p<0.01, ** p<0.05, * p<0.1. Each column is one model as in equation (1) (Panel A) or equation (2) (Panel B) using the according subsample for each specialty of Sample A (Panel A) or Sample B (Panel B). Both models include day-of-week FE, week-of-year FE, calling time of day and practice FE. The coefficient for column (1) in Panel A becomes significant when day-of-week FE are excluded. The sum of the subsamples is slightly smaller than Sample A and Sample B in the main analysis as those samples also include pre-test-observations for the other indications (eye examination, a magnet-resonance-therapy of the right knee, and a pulmonary function test).

Table 6: Effect Heterogeneity

	(1) Physician density	(2) Resident per km ²	(3) Household income	(4) East Germany	(5) Share of PHI in state
Panel A: <i>apptm</i>					
Privately Insured					
*[column]	-0.0022 (0.0120)	-0.0000 (0.0000)	-0.0000 (0.0001)	-0.0093 (0.0495)	0.0016 (0.0028)
Privately Insured	0.0482 (0.0538)	0.0475 (0.0292)	0.0554 (0.1054)	0.0396** (0.0173)	0.0191 (0.0348)
Column	0.0232 (0.0313)	-0.0000 (0.0000)	0.0002 (0.0003)	0.0047 (0.1415)	0.0022 (0.0176)
Panel B: <i>Log(dayswait)</i>					
Privately Insured					
*[column]	-0.1209 (0.0982)	-0.0002** (0.0001)	-0.0004 (0.0004)	0.3615 (0.3957)	0.0161 (0.0296)
Privately Insured	-0.6148 (0.4444)	-0.7679*** (0.1974)	-0.3541 (0.7664)	-1.1814*** (0.1356)	-1.3296*** (0.3792)
Column	-0.2723 (0.2286)	0.0014 (0.0010)	-0.0024 (0.0019)	3.7602*** (0.6305)	0.7537*** (0.1366)
Notes: Robust standard errors in parentheses, clustered at the practice level *** p<0.01, ** p<0.05, * p<0.1. Each column in each panel is one model. Panel A runs models as in equation (1) with 1,328 observations. Panel B runs models as in equation (2) with 1,004 observations. The column header indicates the stratifying variable and [column] represents these variables. The physician density is based on the average physician density of the three specialist groups.					

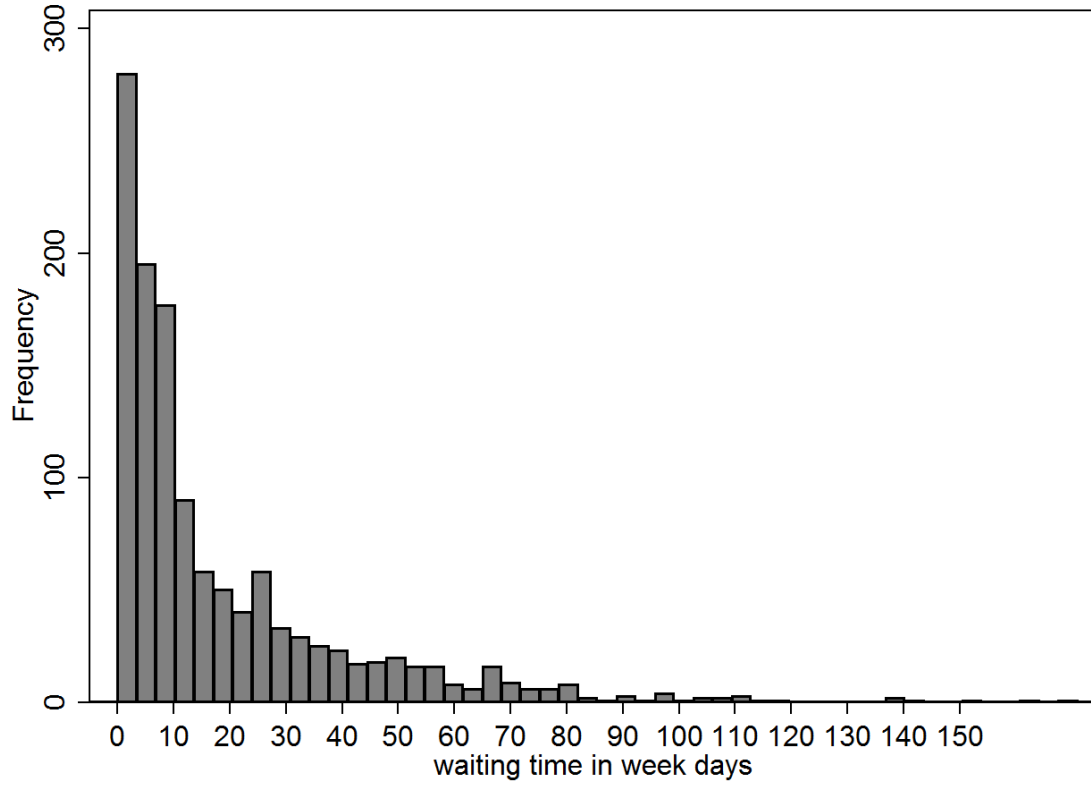
Appendix A

Figure A1: Distribution of Number of Contacted Specialists by 36 Counties



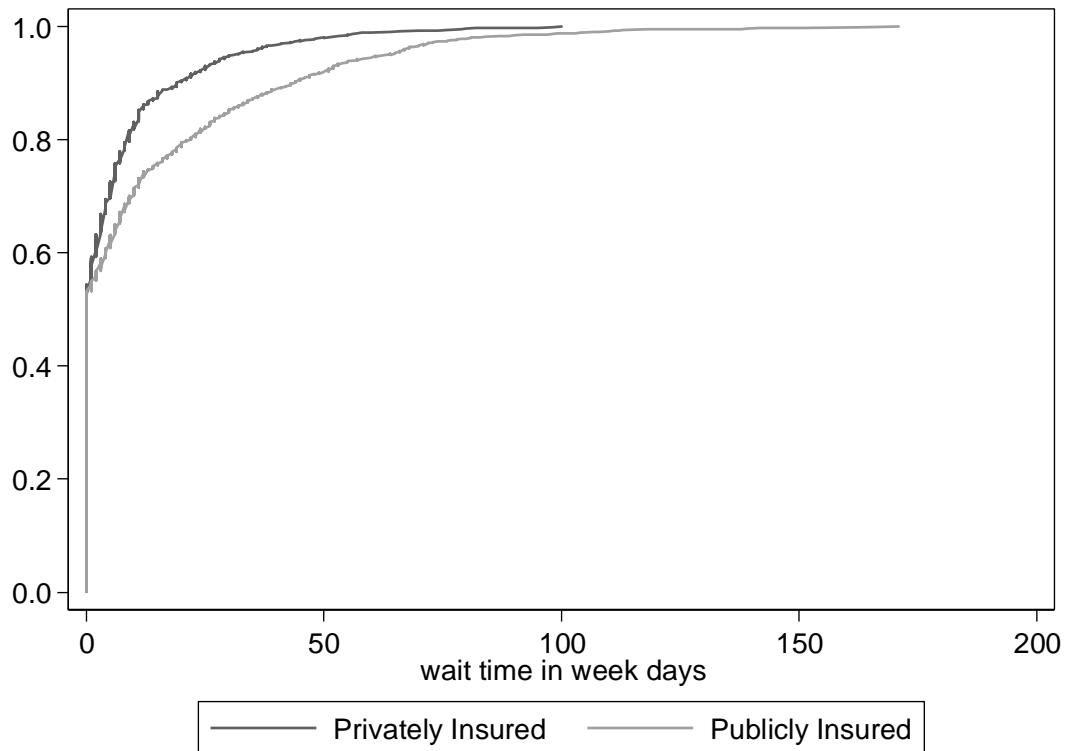
Notes: The histogram displays the number of contacted specialists by county. In total, 36 representative counties were included in the experiment (see Section 4).

Figure A2: Distribution of Wait Time in Weekdays



Notes: The histogram displays the wait time in weekdays for all successfully contacted practices that offered an appointment, i.e., the 81% of practices in Sample A.

Figure A3: Distribution of Wait Time in Weekdays



Notes: The cumulative density functions display the wait time in weekdays for all successfully contacted practices that offered an appointment under both insurance types (i.e. Sample B). The cdf of privately insured patients dominates the cdf of publicly insured patients, i.e. for each wait time in weekdays the cumulative density for privately insured is higher than the density for publicly insured.

Figure A4: Reasons for Exclusion of Practices

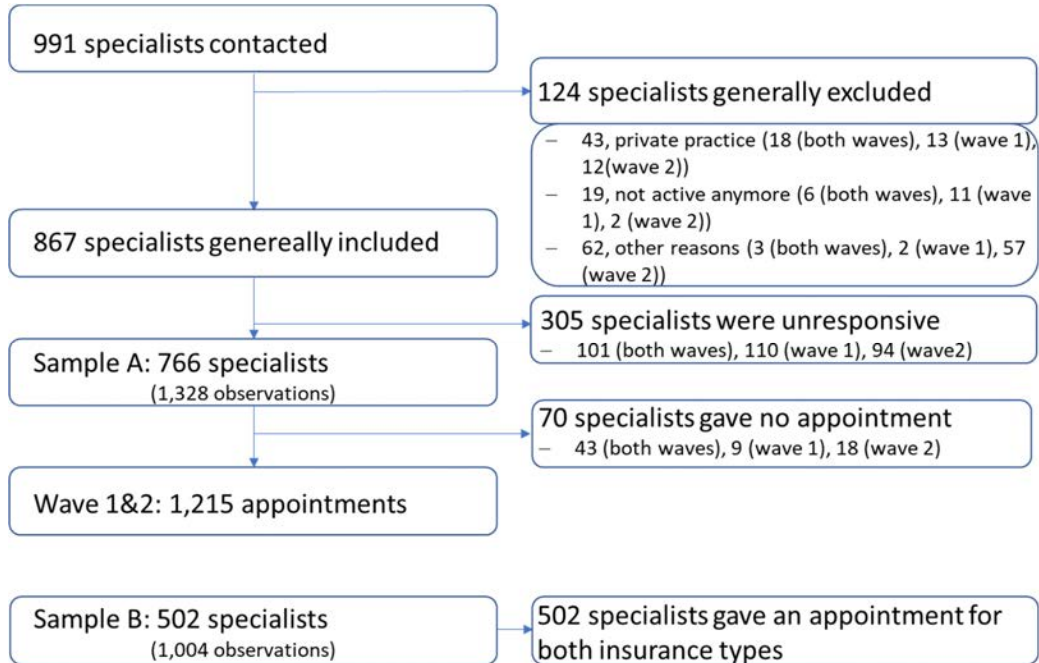


Table A1: Comparison of Publicly and Privately Insured in Germany

	Publicly Insured	Privately Insured	Privately Insured			
			Civil Servants	High income	Self Employed	Non-Employed
Age	50.29	53.69	45.39	48.46	49.50	64.56
Female	0.5254	0.3912	0.4976	0.2431	0.2095	0.4523
Smoker	0.2691	0.1678	0.1852	0.1516	0.2530	0.1316
BMI	26.56	25.86	25.69	25.97	26.11	25.86
Hospital stay last year	0.1369	0.1130	0.1005	0.0636	0.0234	0.1889
Hospital nights last year	1.4380	1.2717	1.3751	0.5016	0.1125	2.0833
Outpatient visits	2.3895	2.3865	2.6044	1.7043	1.3431	3.0272
Public sector employee	0.1126	0.3018	0.9782	0.1616	0.0091	0.0011
Full-time employed	0.3686	0.5079	0.8234	0.8247	0.8736	0.0000
Part-time employed	0.1880	0.1145	0.1744	0.1753	0.1264	0.0000
Dropout	0.0255	0.0078	0.0053	0.0050	0.0242	0.0039
High school degree	0.1830	0.5265	0.6948	0.5821	0.4114	0.4360
Monthly gross wage	2,403	4,708	3,833	6,059	5,341	4,039
Monthly net wage	1,564	3,118	2,922	3,769	3,140	3,029
Equalized HH income	23,228	40,031	34,264	50,957	52,992	34,707
<i>Observations</i>	<i>23,970</i>	<i>3216</i>	<i>773</i>	<i>460</i>	<i>457</i>	<i>1034</i>

Notes: SOEP (2016) v.33 -- 95% sample. See Goebel et al. 2019 for details. All summary statistics are weighted using SOEP cross sectional weights. The number of observations indicated in the last row is smaller for the following variables: smoker, BMI, outpatient visits, monthly gross and net wage as well as the equalized household income.

Table A2: Descriptive Statistics, Practices that responded Once vs Twice

Variables	Responded once		Responded twice	
	(1) mean	(2) sd	(3) mean	(4) sd
Dependent Variables				
Appointment	1	0	1	0
Waiting time in working days	22.80	25.80	17.28	21.42
Waiting time in working days w/o bridging days	22.06	25.36	16.70	20.80
Main Independent Variables				
Privately Insured	0.578	0.495	0.492	0.500
Allergy test	0.313	0.465	0.179	0.384
Hearing test	0.242	0.429	0.400	0.490
Gastroscopy	0.242	0.429	0.141	0.349
First call privately insured	0.493	0.501	0.492	0.500
County-Level Controls				
Unemployment rate in %	8.431	2.744	7.744	2.593
Share employees w/ academic degree in %	9.614	3.472	10.61	3.983
Household income per capita, €	1,733	321.2	1,807	316.0
Residents per doctor	2.213	0.538	2.266	0.589
Residents per square km ² in 1000	1,968	1,481	2,251	1,605
East German state	0.194	0.397	0.112	0.315
Share PHI ¹ in %	12.64	4.486	11.91	4.978
Number of practices:	211	211	502	502
Sources: See Section 4 and 5 for details. County-level controls are taken from BBSR (2018) and Destatis (2018a, b). ¹ Available at the state level. Columns (1) and (2) report the mean and standard deviation (sd) for practices that only responded once and columns (3) and (4) report the mean and standard deviation for practices that responded twice.				

Table A3: Impact of Insurance Status on Likelihood to be Offered Appointment (Probit & OLS)

	(1) Probit apptm	(2) OLS apptm
Privately Insured	0.0542*** (0.0153)	0.0483*** (0.0117)
Day-of-week FE	X	X
Time of day	X	X
Specialty controls	X	X
Observations	1,328	1,328
R-squared		0.0225

Notes: Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Each column in each panel is one model using Sample B. Column (1) reports the marginal effects from the probit model, column (2) the results from the main analysis in the same specification for comparison.

Table A4: Impact of Insurance Status on Wait Times (Poisson & Negative Binomial)

	(1) <i>dayswait</i>	(2) <i>dayswait</i>	(3) <i>dayswait</i>	(4) <i>dayswait</i>
<i>Poisson model</i>				
Privately Insured	-0.7194*** (0.0595)	-0.7262*** (0.0586)	-0.7435*** (0.0579)	-0.7570*** (0.0592)
Day-of-week FE	X	X	X	X
Week-of-year FE	X	X	X	X
Time of day	X	X	X	X
Specialty controls		X	X	X
Practice FE				X
County FE			X	
Observations	1,004	1,004	1,004	1,004
Pseudo R	0.151	0.276	0.353	0.693
<i>Negative binomial model</i>				
Privately Insured	-0.7534*** (0.0574)	-0.7064*** (0.0555)	-0.7464*** (0.0519)	-0.7932*** (0.0549)
Day-of-week FE	X	X	X	X
Week-of-year FE	X	X	X	X
Time of day	X	X	X	X
Specialty controls		X	X	X
Practice FE				X
County FE			X	
Observations	1,004	1,004	1,004	1,004
Pseudo R	0.026	0.045	0.062	0.209

Notes: Robust standard errors in parentheses, clustered on practice level. *** p<0.01, ** p<0.05, * p<0.1. Each column is one model as in equation (2) using Sample B, see Section 5 for details. The mean of the dependent variable *dayswait* indicating the number of weekdays until the offered appointment is 24.89 for publicly and 11.57 days privately insured patients. The overall mean is 18.23 (all values for Sample B, not shown in Table 1). Panel A reports results from the poisson model and panel B reports results from the negative binomial model.

Table A5: Impact of Insurance Status on Wait Times

	(1) <i>dayswait</i>	(2) <i>dayswait</i>	(3) <i>dayswait</i>	(4) <i>dayswait</i>
Privately Insured	-12.5262*** (1.1075)	-12.6861*** (1.1080)	-12.9848*** (1.1385)	-13.1197*** (1.6333)
Day-of-week FE	X	X	X	X
Week-of-year FE	X	X	X	X
Time of day	X	X	X	X
Specialty controls		X	X	X
Practice FE				X
County FE			X	
Observations	1,004	1,004	1,004	1,004
R-squared	0.1314	0.2579	0.3448	0.7213

Notes: Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Each column is one model as in equation (2) using Sample B, see Section 5 for details. The mean of the dependent variable *dayswait* indicating the number of weekdays until the offered appointment is 24.89 for publicly and 11.57 days privately insured patients. The overall mean is 18.23 (all values for Sample B, not shown in Table 1). In contrast to Table 4, this table does not take the logarithm of the dependent variable.

Table A6: Impact of Insurance Status, First Contact + Main Specialties

	(1) apptm	(2) apptm	(3) Log(dayswait)	(4) Log(dayswait)
Panel A: Only first successful contact				
Privately Insured	0.0751*** (0.0180)	0.0731*** (0.0182)	-0.9327*** (0.1192)	-0.9682*** (0.1235)
Day-of-week FE	X	X	X	X
Week-of-year FE	X	X	X	X
Time of day	X	X	X	X
Speciality controls		X		X
County fixed effects		X		X
Observations	766	766	699	699
R-squared	0.2428	0.3448	0.1784	0.2887
Panel B: Only three main specialties				
Privately Insured	0.0396*** (0.0133)	0.0361* (0.0199)	-1.0148*** (0.1049)	-1.0491*** (0.1547)
Day-of-week FE	X	X	X	X
Week-of-year FE	X	X	X	X
Time of day	X	X	X	X
Specialty controls		X		X
County fixed effects		X		X
Observations	979	979	724	724
R-squared	0.0176	0.8689	0.1536	0.7185

Notes: Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Each column in each panel is one model. Columns (1) and (2) report results for *apptm* and columns (3) and (4) report results for *dayswait*. Panel A restricts the sample to solely the first successful contact and Panel B restricts the sample to the three main specialties.

Table A7: Impact of Insurance Status on Wait Times (unbalanced panel)

	(1) Log(dayswait)	(2) Log(dayswait)	(4) Log(dayswait)
Privately Insured	-1.0110*** (0.0907)	-1.0622*** (0.0879)	-1.0786*** (0.0879)
Day-of-week FE	X	X	X
Week-of-year FE	X	X	X
Time of day	X	X	X
Specialty controls		X	X
County FE			X
Observations	1,215	1,215	1,215
R-squared	0.1612	0.2414	0.2964

Notes: Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Each column in each panel is one model as in Table 4 but using Sample A, the unbalanced panel. See main text for details.

Table A8: Effect heterogeneity (only first successful contact)

	(1) Physician density	(2) Resident per km2	(3) Household income	(4) East Germany	(5) Share of PHI in state
Panel A: <i>apptm</i>					
Privately Insured *[column]	-0.0079 (0.0160)	-0.0000 (0.0000)	0.0000 (0.0001)	-0.0121 (0.0659)	-0.0005 (0.0036)
Privately Insured	0.1156 (0.0748)	0.0994** (0.0387)	0.0592 (0.1110)	0.0786*** (0.0210)	0.0872* (0.0491)
Column	0.0027 (0.0144)	0.0000* (0.0000)	0.0000 (0.0001)	-0.0495 (0.0550)	-0.0001 (0.0037)
Panel B: <i>Log(dayswait)</i>					
Privately Insured *[column]	0.0544 (0.0874)	-0.0001 (0.0001)	-0.0000 (0.0004)	0.1811 (0.3906)	0.0083 (0.0250)
Privately Insured	-1.1904*** (0.4044)	- (0.1829)	-0.9501 (0.7245)	-0.9784*** (0.1236)	-1.0606*** (0.3208)
Column	0.0121 (0.0503)	-0.0001 (0.0001)	-0.0003 (0.0002)	0.2755 (0.1840)	0.0227 (0.0176)
Notes: Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Each column in each panel is one model. Panel A runs models as in equation (1) with only first successful contact with 766 observations. Panel B runs models as in equation (2) with only first successful contact with 699 observations. The column header indicates the stratifying variable and [column] represents these variables. The physician density is based on the average physician density of the three specialist groups.					

Appendix B

Selection of Treatment Counties

We selected the 36 treatment counties using the following procedure based on official data from (BBSR, 2018; Destatis, 2018a, b):

1. Within the 16 federal German states, we chose the number of counties to include based on the population and the geographic size of the counties, such that at least one but at most four counties per federal state were included.
2. We ranked all 16 states by their population and their area in km². Then, we built four categories based on these two rankings. The four categories then determined whether we included 1, 2, 3, or 4 counties of this state in the field experiment. For example, Bavaria is the largest German state in terms of size (70,542 km² or 27,236 miles²). It is the second largest German state in terms of its population (12,930,751 residents in 2017). Hence, we included four Bavarian counties in the experiment.
3. Within a state, we then selected counties based on the average household income. First, we assigned all counties to one of five income categories.³⁵ Then, we counted the number of counties in each of the five income categories. For example, Bavaria is a very prosperous state. None of the 70 counties is in the lowest income category, 6 are in the second lowest, 13 in the third lowest, 26 in the second highest and 25 in the highest. Because (2) determined to choose four Bavarian counties, we selected one from each income category. As another example, Brandenburg (a state in East Germany) is not very populous and prosperous. It has 15 counties in the lowest income category, 2 in the second lowest and 1 in the third lowest. Because (2) determined to choose only one county from Brandenburg due to the relatively low number of residents (2,494,648 in 2017), we included a county from the lowest income category.
4. In the last step, we randomly selected the specific county to be included within the income category. For example, steps (2) and (3) determined to choose one of the 15 Brandenburg counties in the lowest income category. We randomly draw this final county. It is gray shaded in Figure 1.

³⁵ (1) €16.274 - €19.148, (2) €19.149 - €20.928, (3) €20.929 - €22.058, (4) €22.059 - €23.443, (5) €23.444 - €25.663.

Appendix C

Experimental Protocol of Telephone Calls

1. *Start of Call*

SHI: Lisa-Marie Müller/Katharina-Sophie Meier is here, hello. I am calling to ask for an appointment. I am insured at Techniker-Krankenkasse and I need an appointment for a gastroscopy/an allergy test/a hearing test.

PHI: Katharina-Sophie Meier/Lisa-Marie Müller is here, hello. I am calling to ask for an appointment. I am privately insured and I need an appointment for a gastroscopy/an allergy test/a hearing test.

2. *End of Call*

Oh, I just realize, I don't have my calendar with me. I need to call again. Don't block the slot for me, please, I just can't say whether I am free then.

3. *Answers to frequently asked questions*

Q: Have you been to our practice before?

A: No, I have never been at your place. I have just moved here.

Q: Why do you need the appointment? Are you in pain?

A: No, it is not urgent. I went to see the GP and he gave me the referral to check it.

Q: Who is your GP?

A: His name is Schmitz, but he is at my home town.

Q: Do you have a referral and what does it say on the referral?

A: Yes, I do. But I don't have it with me right now.

Q: When is your birthday?

A: My birthday is on 20/12/1988

Q: What is your phone number?

Provision of a phone number that was not in use anymore. The actual number of the phone the call was conducted with was suppressed.