RUHR ECONOMIC PAPERS

Anna Werbeck
Ansgar Wübker
Nicolas R. Ziebarth

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

#846

Imprint

Ruhr Economic Papers

Published by

RWI – Leibniz-Institut für Wirtschaftsforschung
Hohenzollernstr. 1-3, 45128 Essen, Germany

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

Editors

Editorial Office

Ruhr Economic Papers #846

Responsible Editor: Ansgar Wübker

ISSN 1864-4872 (online) – ISBN 978-3-86788-981-0

Ruhr Economic Papers #846

Anna Werbeck, Ansgar Wübker, and Nicolas R. Ziebarth

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

Bibliografische Informationen der Deutschen Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at **http://dnb.dnb.de**

RWI is funded by the Federal Government and the federal state of North Rhine-Westphalia.

http://dx.doi.org/10.4419/86788981
ISSN 1864-4872 (online)
ISBN 978-3-86788-981-0

Anna Werbeck, Ansgar Wübker, and Nicolas R. Ziebarth[1]

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

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 more than twice as high 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 7% 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. Our findings show that structural differences in reimbursement rates lead to structural differences in health care access.

JEL-Code: I14, I11, I18

Keywords: Health care inequality; reimbursement rates; health care access; discrimination; cherry picking; gastroscopy; audiometry; allergy test; allergists; otorhinolaryngologist; gastroenterologist

April 2020

1 Anna Werbeck, RGS Econ and RWI; Ansgar Wübker, RWI, LSCR and RUB; Nicolas R. Ziebarth, Cornell University - *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, Daniel Herrera, Tom Hoe, Toshiaki Iizuka, Nima Jouchaghani, Hendrik Jürges, Christoph Kronenberg, Isabelle Maripier, Giuseppe Moscelli, Julien Mousquès, Catia Nicodemo, Cheti Nicoletti, Simon Reif, Hendrik Schmitz, Hannes Schwandt, Erin Strumpf, David Slusky, Chris Whaley, Daniel Wiesen, Lilia Zhurakovska and participants at the 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 Haering and Paul Pangburn for excellent research assistance. 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. -* *All correspondence to: Nicolas Ziebarth, Cornell University, Policy Analysis and Management (PAM), 426 Kennedy Hall, Ithaca, NY 14853, DIW Berlin, e-mail: 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; Sicilinani 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 12 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 12 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 inherent 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 and sick (Reinhardt 2011). However, although plenty of anecdotal and descriptive evidence exists, it is difficult to show in a causal framework that health care providers discriminate against Medicaid enrollees and 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, 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 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. This 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 to be offered an appointment is a highly significant 7% 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.

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 Unites 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 bump[1] 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 suggest that this was likely the case (Alexander and Schnell 2017; Maclean et al. 2018).

[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.

This study is one of the first real-world studies that leverages a large randomized field experiment to test if insurance status causally affects health care access. One single test person called almost one thousand providers twice where we randomized 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—just as in the German case—healthier and wealthier 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 inequality 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]

The next section describes the literature on this topic, followed by a discussion of the institutional setting in Germany. Section 4 explains the setup of our field experiment and Section

[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).

5 the data. After that, we outline the statistical approach of this study before discussing the findings. Section 8 concludes.

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 thin.

In contrast, 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.

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, for example, Brosig-Koch et al. (2017) for lab experiments in Germany. 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 treatments, 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 in addition to the theoretical analyses of Siciliani (2006), Gravelle and Siciliani (2008), and Felder (2008). 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. In the U.S., the means-tested state-level program Medicaid covers the poorest members of society (which also are disproportionately sick). Medicaid pays by far the lowest reimbursement rates of all insurance systems. Several papers have studied the association 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 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.

To our knowledge, there exist three studies (two outside the field of economics) which are similar in design to ours and called providers at least twice with the insurance status randomized. First, between 2002 and 2003, Asplin et al. (2005) called around 500 ambulatory clinics in 9 U.S. cities twice and randomized the insurances status of the caller. They find that a higher share of privately insured patients was offered an urgent ambulatory follow-up visit *within a week* (i.e. they only requested appointments within a week). Second, Kuchinke et al. (2009) scheduled appointments at around 500 acute care hospitals in Germany. They find that privately insured callers are offered appointments 1.6 days faster than publicly insured callers.[3] However, differences in wait times were only estimated conditional on the hospital *inquiring* about the insurance status (only 25% did).[4] Moreover, while private insurance may cover more generous (or different) treatments for privately insured, reimbursement rates do not vary between public and private insurance for inpatient care in Germany. Third, Heinrich et al. (2018) called 163 specialists and evaluated a 2015 reform that intended to reduce wait times for the publicly insured in Germany. They compare data from 2014 to data from 2016, but do not find evidence that the reform reduced wait time differences.

In contrast to these studies, in our setting, a test person deliberatly called each practice *twice* following a standardized protocol, where the insurance status of the caller was randomized. Moreover, the *same person* called all private specialist practices and always 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. Non-urgent settings are those where most patient-provider contacts occur. In addition, our

[3] In a third study outside of economics, Lüngen at al. (2008) also find that privately insured patients were offered appointments faster than publicly insured patients. However, they only called each practice *once* and the "inclusion rates" differed by insurance status; the authors do not show whether practice covariates were balanced and whether the randomization was successful. Two other field experiments also contacted practices only once to 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 employee of a major bank vs. a welfare recipient increases the likelihood of getting an appointment. The latter requested 1,310 appointments by email. They find that closing with "Dr." increases the likelihood of an appointment. Another study outside of economics from the U.S. leveraged experimental data to assess the associations between primary care clinics' provider mix and their accessibility to prospective new patients. Clinics with more non-physician clinicians were associated with better access for Medicaid patients (Richard and Polsky 2016).

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

randomized experiment uses a contemporaneous setting and took place over the course of one year between 2017 and 2018. Finally, 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. As mentioned, we focus on elective care among outpatient specialists. These treatments reflect the regular day-to-day access barriers to health care much better than studying medical emergencies.

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.

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 (*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.[5]

Similar to Medicare, by defining annual values per point, 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)[6]. 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.[7] In comparison, for the state-level Medicaid insurance for low-income populations in the United States, 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

[5] 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).

[6] Geographic adjustment factors take differences in regional living costs into account.

[7] Interestingly, the reimbursement rates for colonoscopy in the U.S. under Medicare are similar. Under CPT code 45380 "Colonoscopy and biopsy", the Medicare feel schedule lists a reimbursement of $212.70 (CMS 2018b).

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 2.3 for personal services. In specific cases, an adjustment factor of 3.5 can be applied for personal services (§5 II-IV GOÄ).[8] 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 difficult. First, the treatments and services listed in each schedule usually do not exactly correspond. Second, the SHI schedule is closer to a bundled payment schedule and reimbursement rates include consultations and follow-up visits. In PHI, physicians typically charge every single service separately under a pure fee-for-service (FFS) schedule. Third, the GOÄ does not include any budget caps. Moreover, the EBM has been constantly updated, whereas the GOÄ has not been updated since 1996.

Walendzik et al. (2008) analyze and compare differences in the reimbursement amounts for the same treatments under SHI and PHI. They compare data from the largest German sickness

[8] 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.

fund with more than 10 million enrollees (*Techniker Krankenkasse*). For the same services, they find that providers charge 2.28 times higher reimbursement rates for privately as compared to publicly insured patients.

Table A2 (Appendix) 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 yield the points and euro values for publicly insured patients, whereas the next two columns yield the points and euro values for the privately insured. While specialists do not have much leeway when charging sickness funds for treatments for the publicly insured, because of the pure FFS schedule and private contracting, specialists have more leeway when charging private insurers for the privately insured.[9] Hence, the last two columns of Table A2 list actual charges for privately insured's treatments when the diagnosis contained the ICD-10 code T78 ("allergy") as well as K29 ("Gastritis and duodenitis"), K30 ("Functional dyspepsia") or K31 "(Other diseases of stomach and duodenum"). We obtained these claims data from one of the largest private German insurers with about half a million enrollees from 2005 to 2011 (Karlsson et al. 2016). As seen, reimbursement rates for allergy tests are three to four times higher for privately insured (€49 vs. €184/201). The absolute and relative price difference for hearing tests is much smaller but rates for the privately insured are still 50% higher (€16 vs. €23). For upper gastrointestinal endoscopies, reimbursement rates for the privately insured are two to three times higher than for the publicly insured (€89 vs. €163/204/244).

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

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

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 select 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 sampled outpatient specialists to schedule appointments using Google Maps along with the websites of the three major German telephone books. To identify 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 treatments 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 as 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.

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. [10] 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 trigger any suspicion about being part of an audit study. Specifically, we randomized 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

[10] 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.

our regression models, and why balance tables would only show variation in (i) to (iii) but in none of the other variables.[11]

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. All calls were made between Monday and Friday during the regular office hours of each practice.[12] During the call, the caller mentioned that she had a referral by her Primary Care Physician (PCP). When asked for the name of the PCP, the caller gave a fictional name and indicated that the practice was in her hometown. 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.

As mentioned, 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.

[11] 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.

[12] 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.

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 offers only treatments for privately insured patients[13] (43, 4.3%), and (iii) other reasons[14] (55, 5.5%). These reasons reduce the number of unique practices in our study by 117 from 991 to 874.

Second, there are other reasons why practices were unresponsive and we could (structurally) not make appointments; e.g. the practice was closed for at least one week, for example during vacations; the practice did not make fixed appointments; the practice temporarily did not accept new patients or the practice was not reachable after several unsuccessful attempts. In cases where we could only schedule one appointment, we only consider the practice once; for example, when the vacations were over. In other words, for all eligible practices that are not entirely excluded due to reasons (i) and (ii) 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 in both cases.[15] We call this unbalanced sample "Sample A;" it has 1,426 observations of successfully contacted practices. Figure A6 shows a sample selection chart and Table 1 shows the descriptive statistic for this full sample.

[13] 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 banned from charging sickness funds, even when demand from private patients is low. We do not consider these practices as relevant to the experiment.

[14] E.g. practices for children only and misleading telephone numbers.

[15] Practices provided several reasons for why no appointment could be offered, some of which may be true and others excuses. For example, a common justification was that the practice would not make fixed appointments or temporarily would not accept new patients. We remain agnostic about why specifically no appointment was offered but test whether, overall, the insurance status had an impact on the likelihood to receive an appointment.

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 practice was willing to schedule an appointment with the fictional patient. As seen in Table 1, in 85% 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.[16] It has only valid values for the 85% of cases when the practice offered an appointment. Figure A2 shows the distribution of *dayswait* and Table 1 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.

[16] This means that 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).

Main Control Variables

The main variable of interest is *privately insured*. Even in the unbalanced Sample A with 1,426 caller-appointment observations, this variable is almost perfectly balanced with 49.5% of all observations representing a privately insured fictional patient (Table 1).

Other important control variables indicate the day of the week, the exact calendar date, the time of the day when the call was made, 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 insurance status conditional on calling a specific practice on a specific day, which implies that controlling for these variables is warranted.

County-Level Control Variables

The final panel in Table 1 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 age of all residents in the 36 counties is 43 years, the average net income is €1,786 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 public and private 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). 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:

$$a_{ip} = \alpha + \beta PHI_i + X_p'\tau + Z_c'\theta + \gamma DOW_{ip} + ʒTOD_{ip} + \delta_t + \rho_p + e_{ip} \tag{1}$$

where a_{ip} stands for our first outcome variable *apptm*, which is binary and indicates—using the unbalanced Sample A—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 to be publicly or privately insured. 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 which are available upon request.)

Our second model uses the balanced Sample B and is:

$$\ln (w_{ip}) = \alpha + \beta PHI_i + X_p'\tau + Z_c'\theta + \gamma DOW_{ip} + \delta_t + \rho_p + e_{ip} \qquad (2)$$

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%). The five percentage point difference is statistically significant at the 5% level. A formal t-test has a t-value of 2.5 and is statistically significant at the 5% 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. This prior is confirmed by a formal t-test which is significant at the 0.1% level with a t-value of 9.8.[17]

Finally, Figures A4 and A5 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 A2), this heterogeneity test may link cream-skimming to actual differences in patient profitability. It is reassuring to see that, on the extensive margin, systematic patient selection only exists for gastroenterologists and allergists where the absolute and relative price differences are the largest. By contrast, we do not find evidence for systematic patient selection for otorhinolaryngologist (hearing tests), where reimbursement rates are much lower and the

[17] 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.

absolute and relative price differences much smaller (€16 vs. 23). Reassuringly, the same pattern also holds for *dayswait*—although we also find significant wait time differences for hearing tests, they are much smaller than for allergy tests and gastroscopies.

Parametric Findings

Next, we move on to our parametric findings and multivariate regression models. Table 2 shows the findings from our first model in equation (1), which uses the binary *apptm* measure as outcome variable. These models assess the impact of the insurance status on the likelihood to be offered an appointment for non-urgent treatments. Each column in Table 2 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 2 about here]

Table 2 shows the following: First, we find that the insurance status of the caller has a highly significant impact on the willingness 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 4.4 to 6.2 percentage points or by about 7% relative to the mean of 0.85. Second, the estimates are robust across all model specifications and far from statistically different from one another. The inclusion of week-of-year fixed effects, county fixed effects, and even practice fixed effects barely alter the size of the coefficients. Third, the findings are also robust to running probit models and calculating marginal effects (available upon request).

Table 3 follows the same setup as Table 2 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 impact of being privately insured on the mean wait time in weekdays, relative to being publicly insured.[18] In the Appendix, in Table A3, we replicate Table 3 but do not take the logarithm of the dependent variable. All six models in Table 3 and Table A3 use our balanced Sample B and only include the 502 unique practices that offered specific appointments to both callers, the publicly and the privately insured.

[Table 3 about here]

Table 3 and Table A3 show the following: All six model coefficients are highly significant at the 1% level. Moreover, the estimates are very robust to the sets of covariates included, reinforcing that our randomization was very successful. It 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 t-tests, it also suggests the absence of structural imbalances by day-of-the-week or the time of the day when the call was made.

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).

Table A4 shows regressions for *apptm* and *dayswait* but separately for the three specialist groups. The estimates in Panel A (*apptm*) all have the expected sign but are statistically not

[18] Note that the models in Table 3 use 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.

significant plausibly due to the smaller sample size. The estimates in Panel B (*dayswait*) are very similar to the main analysis and statistically significant for each specialty. In terms of effect size, although we do find significant waiting time differences for hearing tests, they are smaller than the effect sizes for allergy tests and gastroendoscopies. This finding is in line with the smaller reimbursement rate differences between PHI and SHI for hearing tests (also see discussion below).

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, 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 4 shows the results for *apptm* and Panel B of Table 4 shows the results for *dayswait*.

[Table 4 about here]

As seen, few of the interaction terms (which indicate differences in insurance status by the stratifying covariate) are statistically significant. The findings for East Germany are relatively large and the sign of the effects consistent with the notion that the differences in East Germany are *smaller* than in West Germany. However, the two interaction terms in column (4) of Table 4 are only significant at the 20% level and rather suggestive. Second, the findings for physician density (column 1), household income (column 3) and share of PHI-patients (column 5) are all far from being significant in both panels. Finally, the 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.

In summary, we find that inequality in wait times is larger when the population density in the county is larger. We also find suggestive evidence that inequality in access is less pronounced in East as compared to West Germany, possibly indicating a long-term effect of socialist norms (Alesina and Fuchs-Schündeln 2007; Rainer and Siedler 2009). For example, in a recent survey, PWC (2017) finds that a 10 percentage point higher share of West as compared to East Germans have a positive attitude towards more competition in the health care sector (59% vs. 49%).

Discussion

Our field experiment has clearly established 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.

As we randomized the insurance status of the fictional patient in our study, our "cream-skimming" interpretation of the findings allows for several specific mechanisms. To be specific, while it is well-known—particularly among medical professionals—that private insurers pay much higher rates, it is also true that the privately insured have higher incomes, are better educated and are healthier, see Table A1 (Appendix). However, we believe that all specific explanations can ultimately be subsumed as: *practices structurally select more profitable patients*.[19]

[19] 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).

First, doctors may expect to not only receive a significantly higher reimbursement from the privately insured but also deal with healthier patients who have fewer co-morbidities. Given a specific diagnosis, healthier patients imply shorter treatments and interactions and more profitable patients, especially when payments are bundled. (A counterargument is that 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 a proxy for more profitable, wealthy patients (and not vice versa). Moreover, there is evidence that doctors consider highly-educated, "empowered" patients as rather annoying, time-consuming and difficult (Neuberger 2000; Rankin 2011).

Third, in Germany, the privately insured have, without any doubt, higher incomes (Table A2) 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 A2, 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, medically unnecessary and not covered services to the more affluent privately insured. On the other hand, public insurance plans have basically zero cost-sharing and public insurers also directly pay providers without intensive claim reviews.[20] Private insurers usually share costs, review claims, and do not pay providers

[20] 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

directly, hence, the risks of claim disputes and non-payments are higher when treating privately insured.[21]

Lastly, as shown in Figures A3 and A4, there is evidence that inequality in access is less pronounced—and not existent at the extensive margin—for hearing tests where reimbursement rates are the lowest and the difference between private and public rates the smallest (Table A2). This finding suggests that the specific reimbursement rate differentials do drive the selection of the more profitable privately insured.

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).[22] In the German case, 77% of the population do not have the option to purchase 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, poverty and age. As mentioned in the introduction, inequality in access to health care is just *one* dimension on which health care systems are rated. Trade-offs with other

---

Council „*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).

[21] Recall that the patient first pays providers and then submits the claim to the insurer.

[22] 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.

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 with classic fee-for-service schemes without caps or bundled payments. Importantly, reimbursement rates for both systems are standardized and centrally set, instead of through individual negotiations between insurers and providers. Provider networks do not exist in Germany and hence do not operate as a confounding factor.

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 seven 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.[23] What's more, most Germans seem to find the structural differences between the two systems unacceptable. In a representative survey among Germans, two thirds indicated 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 physicians or increase taxes or premiums. How optimal

[23] 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). What's more, if a person receives treatment before the stomach cancer spreads, the 5-year survival rate is 68 %; once it reaches distant organs, survival rate drops to 5%.

reimbursement rates should be set in a health care system in practice is an important topic for future research.

References

Alesina, A., Fuchs-Schündeln, N. 2007. Goodbye Lenin (or not?): The effect of communism on people's preferences, *American Economic Review*, 97 (4): 1507-1528.

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)

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 Medicaire and Medicaid Services 2018a. *Physician Fee Schedule,* https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeeSched/index.html, (date of access: September 10, 2019).

CMS – Centers for Medicaire and Medicaid Services 2018b. CY 2018 *Medicare Fee Schedule Changes*, 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, 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, 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.

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*. 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-Würtemberg Mosbach. 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. https://www.gkv-spitzenverband.de/english/statutory_health_insurance/statutory_health_insurance.jsp, (date of access: September 10, 2018).

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. 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,* 238 (5): 375-394.

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.

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

KBV - Kassenärztliche Bundesvereinigung 2017. *Einheitlicher Bewertungsmaßstab (EBM).* 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*. http://www.kbv.de/html/2054.php (date of access: September 10, 2018).

KBV - Kassenärztliche Bundesvereinigung 2018b. *Einheitlicher Bewertungsmaßstab (EBM).* 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.* 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.

Lindsay, C.M., Feigenbaum, B. 1984. Rationing by waiting lists, *The American Economic Review*, 74(3), pp. 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., Ãeter, 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.

PWC 2017. *Bevölkerungsbefragung: Wuensche and das Gesundheitssystem,* September 2017. https://www.pwc.de/de/gesundheitswesen-und-pharma/berichtsband-gesundheitssystem.pdf (date of access: October 10, 2018).

Rainer, H. Siedler, T. 2019. Does democracy foster trust? *Journal of Comparative Economics*, 37(2): 251-269.

Rankin, L. 2011. When Docs Get Annoyed At Empowered Patients, Psychology Today, Jul 19, 2011. https://www.psychologytoday.com/us/blog/owning-pink/201107/when-docs-get-annoyed-empowered-patients, (date of access: July 7, 2019).

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.

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., 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.

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.

Soester Anzeiger 2019. Warum Patienten lange auf einen Facharzt-Termin warten, https://www.soester-anzeiger.de/lokales/soest/manche-fachaerzte-lehnen-neue-patienten-12927558.html, (date of access: March 13, 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*. 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.

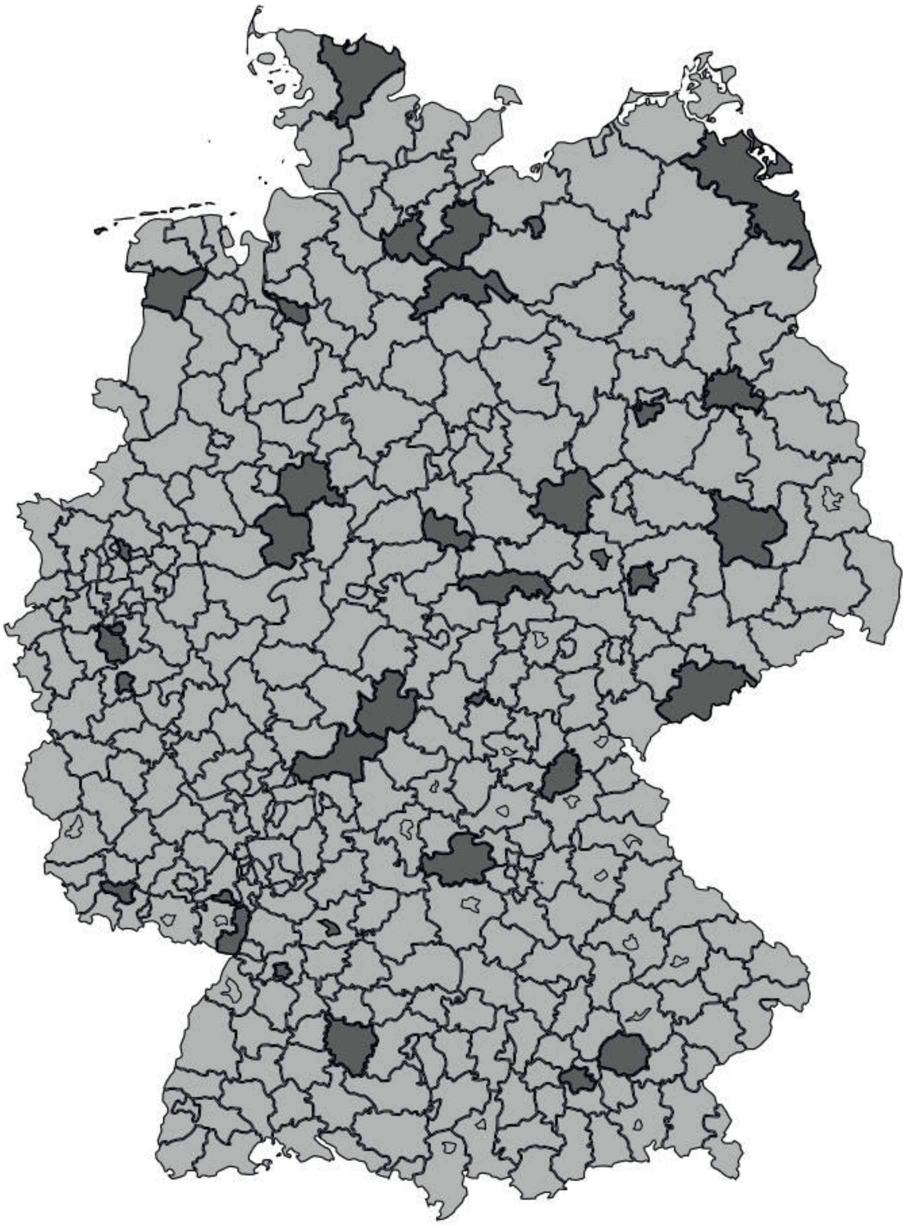
Walendzik, A. und Wasem, J. (2019), *Ärztliche Vergütung ambulanter und ambulant erbringbarer Leistungen Gesundheitspolitisch zielgerechte Integrationsmodelle über sektorale Leistungsträger und Finanzierungssysteme,* Bertelsmann Stiftung (Eds.). First edition, https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/aerztliche-verguetung-ambulanter-u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

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.

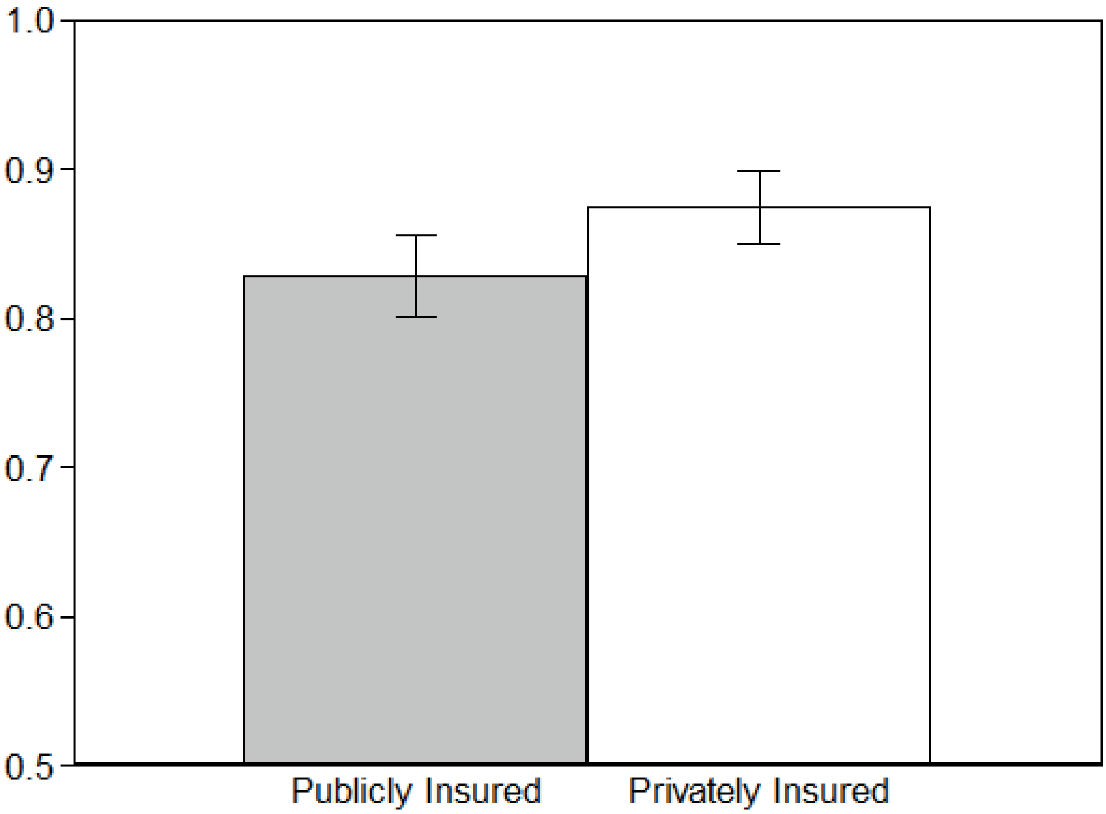
Figures and Tables

Figure 1: Selected Counties for Field Experiment



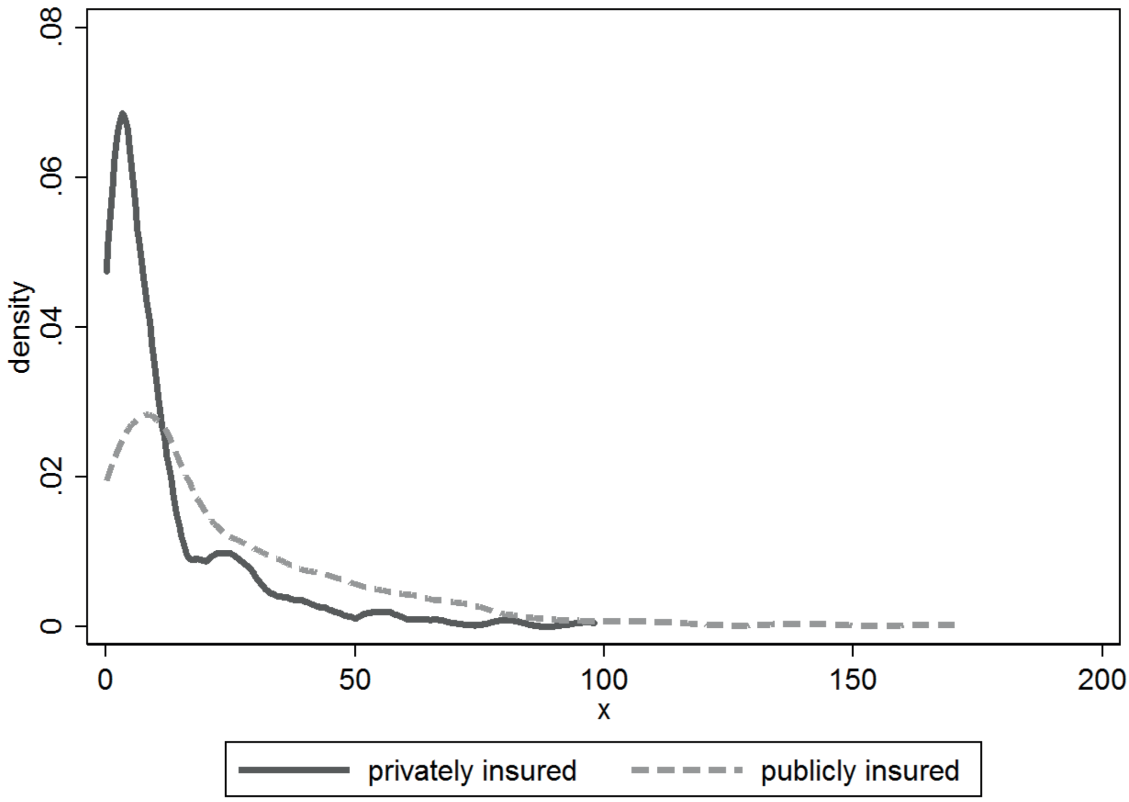
Source: Own illustration. The 36 counties selected for the field experiment are dark gray. Between one and four counties in each of the 16 federal states were selected (see Appendix B).

Figure 2: Likelihood to be Offered Appointment by Insurance Status



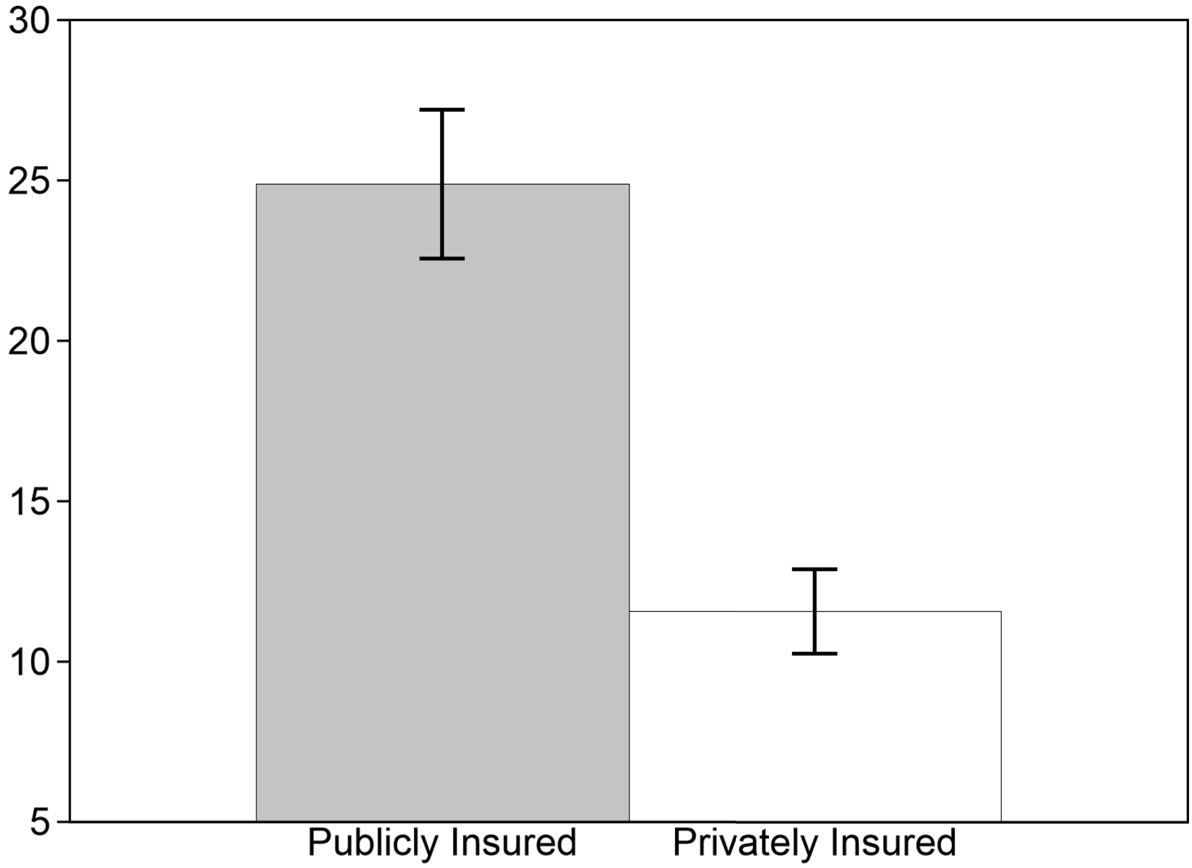
Source: Graph uses Sample A. The bars show 95% confidence intervals.

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.

Table 1: Descriptive Statistics

VARIABLES	(1) mean	(2) sd	(3) min	(4) max	(5) N
Dependent Variables					
Apptm	0.852	0.355	0	1	1,426
Dayswait	19.05	23.16	0	171	1,215
Dayswait II	18.36	22.73	0	171	1,215
Main Independent Variables					
Privately Insured	0.500	0.500	0	1	1,426
Allergy test	0.219	0.414	0	1	1,426
Hearing test	0.353	0.478	0	1	1,426
Gastroscopy	0.171	0.377	0	1	1,426
First call privately insured	0.492	0.500	0	1	1,426
County-Level Controls					
Unemployment rate in %	7.947	2.654	2.100	13.60	1,426
Share Employees w/ academic degree in %	10.31	3.861	3.800	16.80	1,426
HH income per capita, €	1,785	319	1,362	3,451	1,426
Physician density	4.35	1.29	1.33	6.83	1,426
Residents per km^2 in 1000	2,167	1,573	55	4,668	1,426
Eastern federal state	0.136	0.343	0	1	1,426
Share PHI[1] in %	12.127	4.844	4.354	17.874	1,426

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

Table 2: Impact of Insurance Status on Likelihood to be Offered Appointment

	(1) *apptm*	(2) *apptm*	(3) *apptm*	(4) *apptm*
Privately Insured	**0.0441**** **(0.0202)**	**0.0541***** **(0.0200)**	**0.0587***** **(0.0204)**	**0.0618**** **(0.0287)**
Day-of-week FE	X	X	X	X
Month-of-year FE	X			
Week-of-year FE		X	X	X
Time of day	X	X	X	X
Specialty controls		X	X	X
Practice FE				X
County FE			X	
Observations	1,426	1,426	1,426	1,426
R-squared	0.0404	0.0887	0.1141	0.4844

Sources: 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 3: Impact of Insurance Status on Wait Times

	(1) *Log(dayswait)*	(2) *Log(dayswait)*	(3) *Log(dayswait)*	(4) *Log(dayswait)*
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
Month-of-year FE	X			
Week-of-year FE		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

Sources: 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 binary variable *daywait* 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.

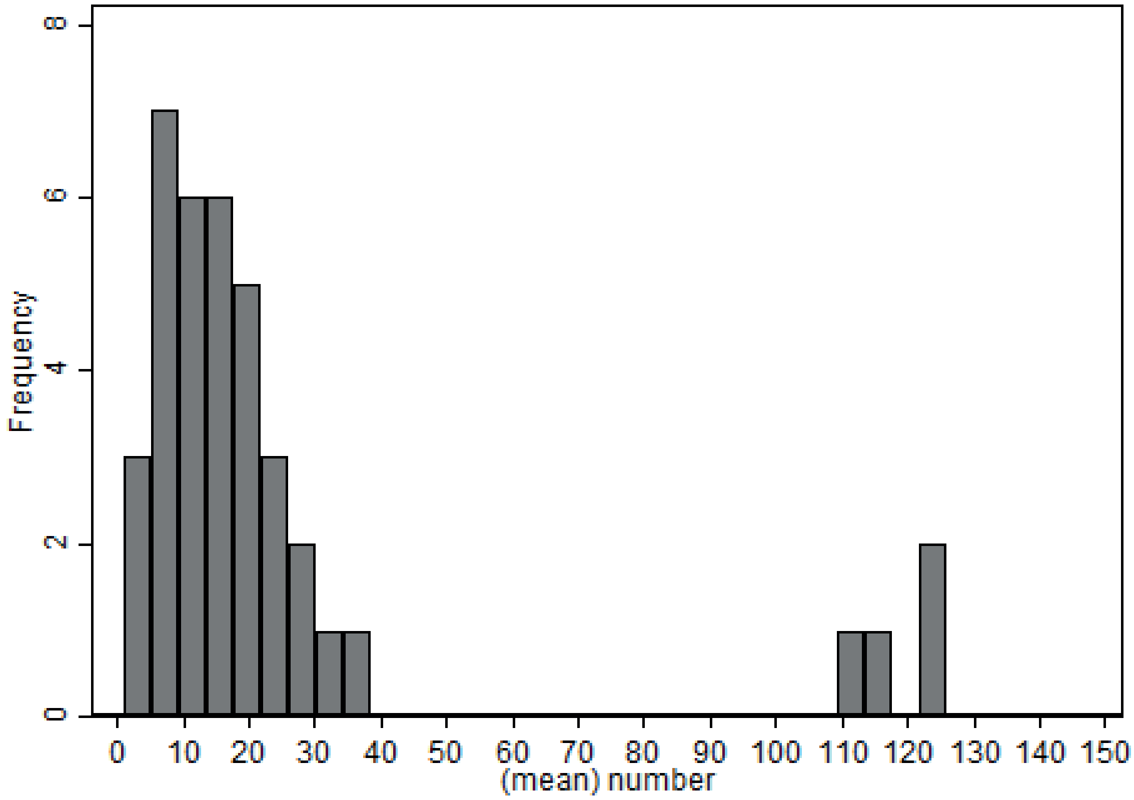
Table 4: Effect Heterogeneity

	(1) Physician density	(2) Resident per km2	(3) Household income	(4) East Germany	(5) Share of PHI in state
Panel A: apptm					
Privately Insured *[column]	**-0.0162**	**-0.0000**	**0.0001**	**-0.1328**	**-0.0007**
	(0.0232)	**(0.0000)**	**(0.0001)**	**(0.0959)**	**(0.0056)**
Privately Insured	0.1320	0.0662	-0.0294	0.0810***	0.0708
	(0.1046)	(0.0498)	(0.1774)	(0.0297)	(0.0715)
Column	0.3110***	-0.0013***	0.0024***	0.3581*	-0.0223
	(0.0286)	(0.0001)	(0.0002)	(0.2073)	(0.0250)
Panel B: Log(dayswait)					
Privately Insured *[column]	**-0.1209**	**-0.0002****	**-0.0004**	**0.3615**	**0.0161**
	(0.0982)	**(0.0001)**	**(0.0004)**	**(0.3957)**	**(0.0296)**
Privately Insured	-0.6148	-0.7679***	-0.3541	-1.1814***	-1.3296***
	(0.4444)	(0.1974)	(0.7664)	(0.1356)	(0.3792)
Column	-0.2723	0.0014	-0.0024	3.7602***	0.7537***
	(0.2286)	(0.0010)	(0.0019)	(0.6305)	(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 1,426 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.

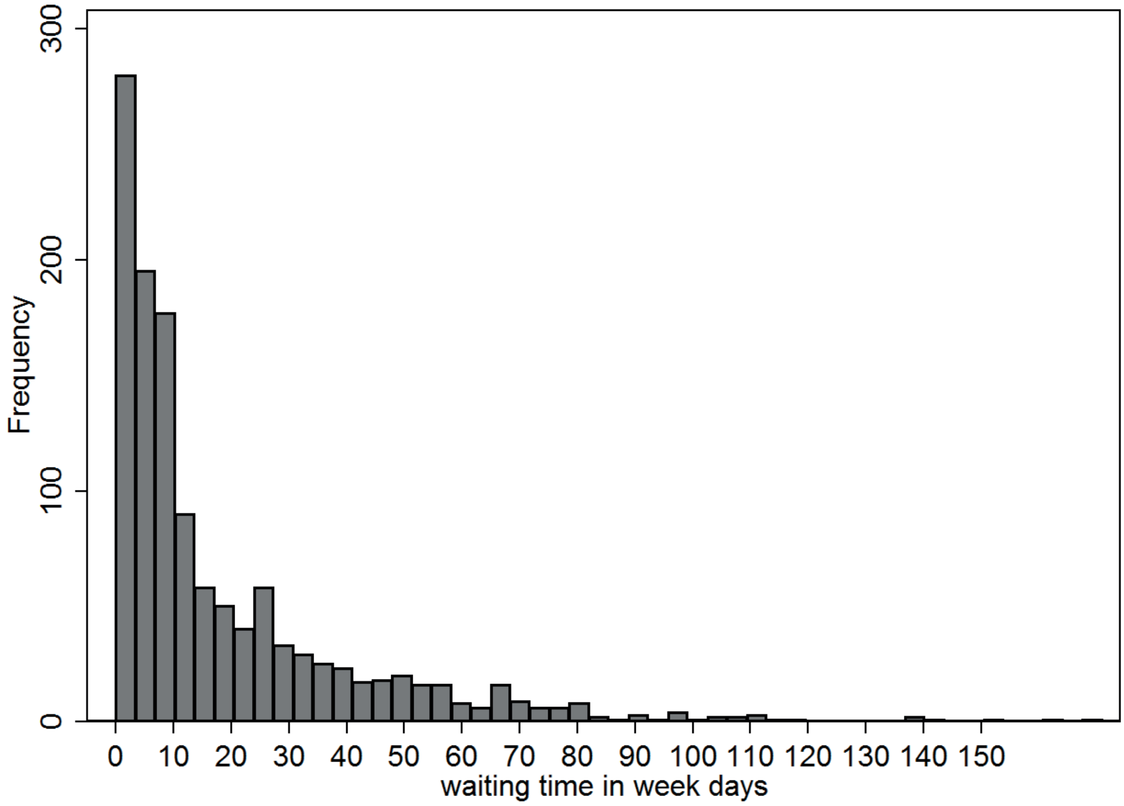
Appendix A

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



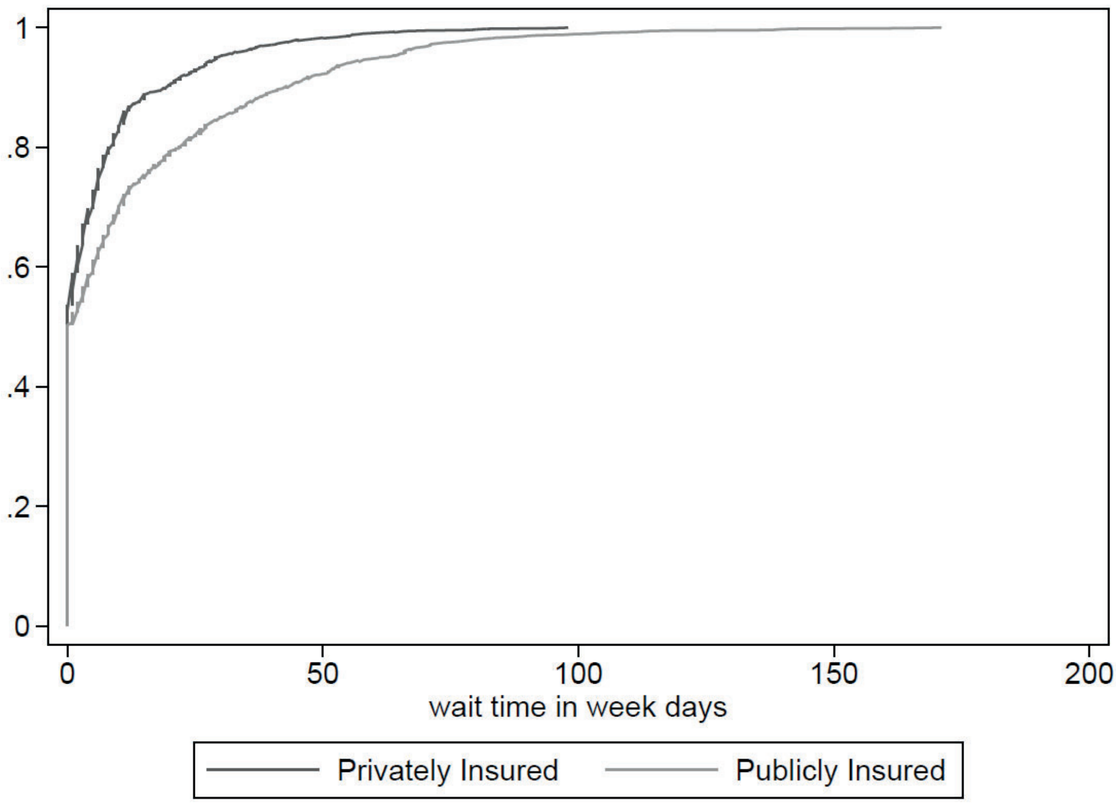
Note: The histogram displays the number of contacted specialists by county. In total, 36 representative counties were included in the experiment (see Section 4).

Figure A2: Distribution of Wait Time in Weekdays



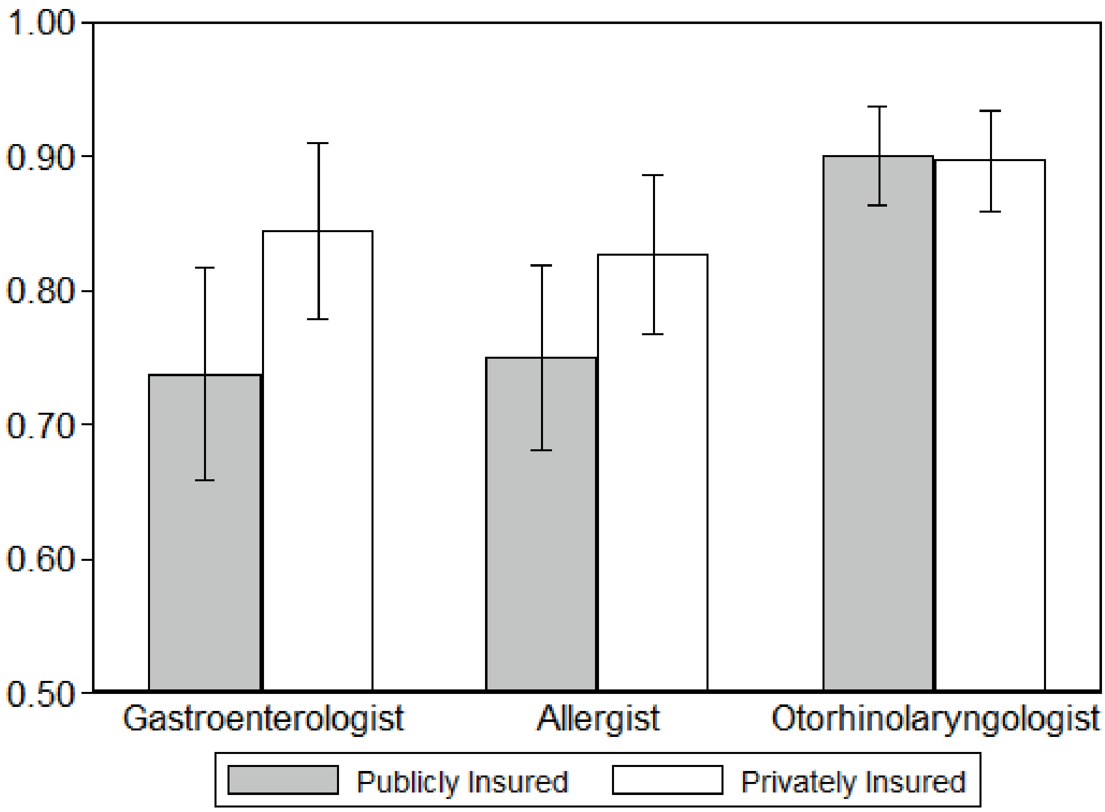
Note: The histogram displays the wait time in weekdays for all successfully contacted practices that offered an appointment, i.e., the 81% of practices in Sample A.

Figure A3: Distribution of Wait Time in Weekdays



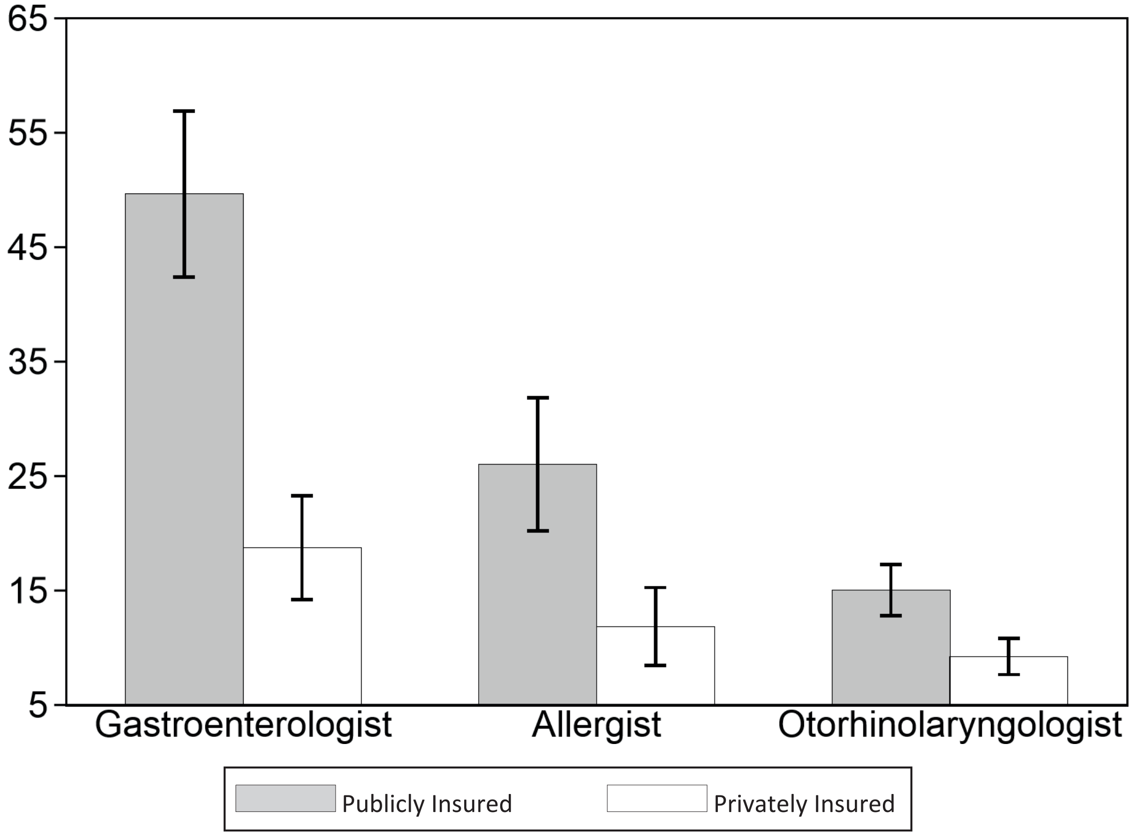
Note: The cumulative density functions display the wait time in weekdays for all successfully contacted practices that offered an appointment under both insurance types (i.e. Sample B). The cdf of privately insured patients dominates the cdf of publicly insured patients, i.e. for each wait time in weekdays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

Figure A4: Likelihood to be Offered Appointment by Insurance Status and Specialist



Source: Graph uses Sample A. The bars show 95% confidence intervals.

Figure A5: Average Wait Times by Insurance Status and Specialist



Source: Graph uses Sample B. The bars show 95% confidence intervals.

Figure A6: Reasons for exclusion of practices

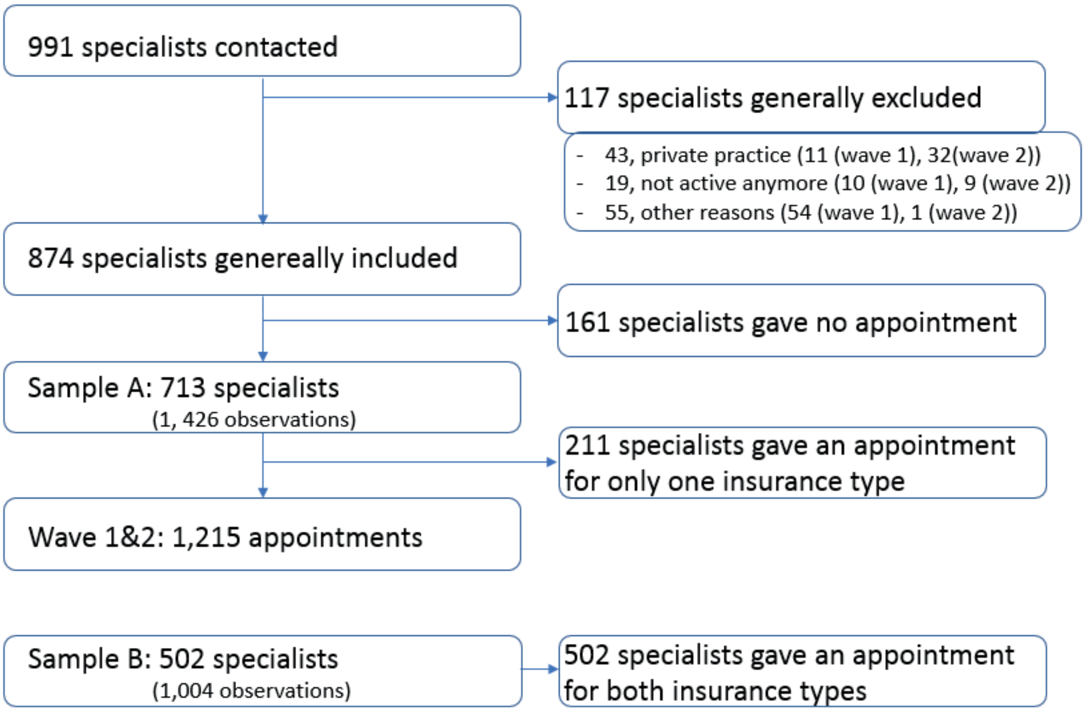


Table A1: Comparison of Publicly and Privately Insured in Germany (SOEP data, 2016)

			Privately Insured			
	Publicly 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
Less accomplished due to physical issues	0.3618	0.2551	0.2177	0.1867	0.1741	0.3459
Daily activities limited due to physical issues	0.3417	0.2303	0.1708	0.1449	0.1427	0.3399
Less accomplished due to emotional issues	0.2204	0.1443	0.1233	0.0986	0.1019	0.1946
Less careful in daily activities due to emotional issues	0.1822	0.1050	0.0889	0.0886	0.0585	0.1413
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
Equivalized HH income	23,228	40,031	34,264	50,957	52,992	34,707
N	*23,970*	*3216*	*773*	*460*	*457*	*1034*

Notes: SOEP v.33 -- 95% sample. 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 equivalized household income.

Table A2: Possible SHI and PHI Reimbursement for Select Treatments (Illustration)

	EBM		GOÄ		
	Points	Euros	Points	Euros	Average claim amount per visit
Allergy test	458	48.80	45	183.60	201.49 (450.24)
Ophthalmology	147	15.66	158	23.02	N/A
Gastroenterology	835	88.96	800	163.20/204.02	244.47 (372.38)

Notes: The first two columns display public reimbursement rates according to the *Einheitlicher Bewertungsmaßstab* (EBM) or "Unified Assessment Scale." The next two columns display private reimbursement rates according to the *Gebührenordnung für Ärzte* (GOÄ) or "Fee Schedule for Physicians." The latter assumes an adjustment factor of 3.5. Values for GÖA allergy tests are calculated 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 are in parenthesis.

Table A3: Impact of Insurance Status on Wait Times

	(1) *dayswait*	(2) *dayswait*	(3) *dayswait*	(4) *dayswait*
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
Month-of-year FE	X			
Week-of-year FE		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 binary 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 3, this table does not take the logarithm of the dependent variable.

Table A4: Impact of Insurance Status on Wait Times

	(1) Gastroscopy	(2) Allergy test	(3) Hearing test
Panel A: apptm			
Privately Insured	**0.1230**	**0.0972**	**0.0154**
	(0.0772)	**(0.0931)**	**(0.0407)**
Observations	244	312	504
R-squared	0.666	0.476	0.558
Panel B: Log(dayswait)			
Privately Insured	**-1.482*** **	**-1.225*** **	**-0.834*** **
	(0.341)	**(0.314)**	**(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 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).

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 km^2. 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 km^2 or 27,236 $miles^2$). 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.[24] 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.

[24] (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.