HEDG

HEALTH, ECONOMETRICS AND DATA GROUP

THE UNIVERSITY of York

WP 13/12

Analyzing Regional Variation in Health Care Utilization Using (Rich) Household Microdata

Peter Eibich and Nicolas R. Ziebarth

April 2013

york.ac.uk/res/herc/hedgwp

Analyzing Regional Variation in Health Care Utilization Using (Rich) Household Microdata*

forthcoming in the Special Issue of *Health Policy:*
"Geographic Variation in health care–40 years of "small-area variation"
http://www.journals.elsevier.com/health-policy/

Peter Eibich[1] and Nicolas R. Ziebarth[2]

April 2013

Abstract

This paper exploits rich SOEP microdata to analyze state-level variation in health care utilization in Germany. Unlike most studies in the field of the Small Area Variation (SAV) literature, our approach allows us to net out a large array of individual-level and state-level factors that may contribute to the geographic variation in health care utilization. The raw data suggest that state-level hospitalization rates vary from 65 percent to 165 percent of the national mean. Ambulatory doctor visits range from 90 percent to 120 percent of the national mean. Interestingly, in the former GDR states doctor visit rates are significantly below the national mean, while hospitalization rates lie above the national mean. The significant state-level differences vanish once we control for individual-level socio-economic characteristics, the respondents' health status, their health behavior as well as supply-side state-level factors.

Keywords: small area variation, health care utilization, SOEP

JEL codes: I12, I14, I18

[1] DIW Berlin, Mohrenstrasse 58, 10117 Berlin, Germany, and University of Hamburg, e-mail: peibich@diw.de.

[2] Corresponding author: Cornell University, Policy Analysis and Management (PAM), 106 Martha van Rensselaer Hall, Ithaca, NY 14853, DIW Berlin, and IZA Bonn, e-mail: nrz2@cornell.edu, Phone: +1-(607) 255-1180, Fax: +1-(607) 255-4071.

* Peter Eibich gratefully acknowledges generous support under "BASEII" by the Bundesministerium für Bildung und Forschung (BMBF; "Federal Ministry of Education and Research", 16SV5537). We would like to thank Adam Lederer for excellence in editing this paper. We take responsibility for all errors in and shortcomings of the article. 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 that might constitute a conflict of interest.

Introduction

In 1973, John Wennberg and Alan Gittelsohn (1973) published their seminal "Small Area Variations in Health Care Delivery" paper in Science [1]. This paper laid the foundation for one of the most fruitful and policy influential multidisciplinary strands of health research: the Small Area Variations (SAV) literature. In the following four decades, in addition to John Wennberg and his co-authors, an incredibly large number of medical scientists, sociologists, epidemiologists, public health scientists, and economists contributed to this field. In principle, all of these studies seek to explain, mostly empirically, why we observe geographic variation in health care utilization and spending. The main puzzling stylized fact of this research field was noted by Wennberg in 1973: there is not only tremendous geographic variation in health care utilization and spending, but this variation seems to be uncorrelated—if not negatively correlated—with health outcomes and quality of care. One intuitively appealing conclusion would be that this correlation demonstrates the potential for efficiency improvements in the health care delivery process. In other words: there seems to be a way to reduce health expenditures without reducing population health or to increase population health without the need to increase health expenditures.

Unfortunately, this conclusion is based on correlations, not on studies proving causation. There could be many reasons why health care utilization and health outcomes are negatively correlated at the aggregated regional level: one obvious reason is that sicker people face greater health care needs and that people's health may vary systematically across regions. Of course, the community of scholars tries to control for health status as a confounding factor. One promising direction of research is to focus on unambiguous and specific medical conditions that imply few treatment options, such as heart attacks or chronic diseases [2, 3]. Studies focusing on very specific diagnoses typically find a significant geographic variation in treatment styles, given the diagnosis [4, 5]. And again, the question pops up whether this is clear evidence of inefficiencies and wasting of resources [6].

One explanation for the observed empirical relationships is supplier-induced demand—the hypothesis is that doctors and hospitals create their own demand by carrying out too many medically unnecessary treatments in order to maximize their profits [7]. It is relatively easy to show that a high provider density is correlated with frequent, technology intensive, and expensive health care treatments [8, 9]. However, this finding does not prove the existence of supplier-induced demand [10]. Empirically, it is extremely challenging to ultimately prove the existence of supplier-induced demand since, for example, one needs to disentangle patient preferences from typically unobserved physician "inducement."

A related explanation for the observed empirical relationships refers to practice styles—the hypothesis is that there exists regional variation in how medical scientists practice, including a large degree of time persistence [11, 12]. Time persistence in practice styles is proven to exist [13]. However, at the same time, it is shown that variation in practice style is greater within than between geographical markets [14].

In addition to supply-side factors, demand-side factors that vary across geographical regions may also play a crucial role in producing the observed correlations, e.g., cultural habits, opportunity costs, income, or preferences [15]. These demand-side factors are typically harder to observe and measure. The literature mainly focuses on supply-side factors.

A fundamental question, although little debated, is the level of (dis)aggregation that one could or should choose for the empirical SAV analysis [16]. Since 1993, the DARTMOUTH ATLAS OF HEALTH CARE has provided detailed descriptive information on how the health care infrastructure, as well as health care utilization and spending, vary at a disaggregated geographical level in the US. Information is available on the level of the 3,000 US counties, the 3,500 Hospital Service Areas (HSA), or the 300 Hospital Referral Regions (HRR) [17]. Research from European countries is often based on the state or even the country level, given the substantially smaller geographic areas [18-20]. On the other hand, recent papers analyze treatment and utilization behavior at a much more disaggregated level, for example, the hospital or practice levels [21, 22].

Another fundamental question refers to the underlying data that SAV analysis is based upon. The majority of the literature comes from the US and makes use of Medicare claims data. Medicare is the public US health insurance system for the elderly, i.e., those above 64 years. Claims data have the great advantage that the researcher observes the full range of claims submitted, i.e., all diagnoses and treatments that were carried out in each regional unit. This allows the researcher to identify very specific diagnoses and conditions. The number of treatments per resident can then be plotted by regional unit to illustrate regional variation. The same holds true for health care spending, since claims data mostly include reimbursement rates. However, spending measures blur the variation of prices and procedures, thus making it difficult to disentangle the separate impact of the two factors [23]. Clearly, observing the universe of diagnoses and treatments is the outstanding advantage of claims data. On the other hand, the obvious disadvantage is the lack of socio-economic background and general health information. Most importantly, the health of the non-treated is not observed and, thus, analyses are inherently truncated.

Recent strands of the SAV literature focus on the geographic variation in the incidence rates of rare diseases or avoidable deaths. Avoidable deaths are deaths due to diseases that are considered as avoidable in case of timely preventive care and appropriate treatment [24-29].

This paper contributes to the literature by addressing several of the issues discussed above: First and foremost, this study does explicitly not focus on claims data. Instead, this is one of the very few papers in the small area variations literature that makes use of rich and representative household panel data. Using this approach, we are able to consider a rich set of mostly demand-side factors that may explain regional differences in health care utilization. While there has been a decent amount of modeling on the supply side, individual-level demand side factors seem to be one key channel that may help explain the observed regional differences in health care usage. The panel dataset we use includes information on the annual number of nights in a hospital as well as the number of outpatient doctor visits. We consider these measures as sufficient indicators for the overall individual use of the health care system as well as for individual health care spending. While we do not observe specific diagnoses or treatments, the great advantage of the data is the richness of the background information that we can exploit. We observe detailed socio-economic background information, such as the labor market status, marital status, and income. Most importantly, we exploit general individual health measures. The data we use include subjective health measures as well as two quasi-objective generic health measures from the SF12. Regional differences in population health are, presumably, the main determinant for regional differences in health care utilization.

A second principle advantage of using individual-level household data is their representativeness. Studies that rely on diagnoses, practice styles, or incidences of diseases are typically also representative, but only for those who use and have access to health care. However, we observe all individual measures for all residents of a geographical unit—not just for those who utilize health care. For example, consider a situation where person X lives in region A and is objectively sicker than person Y in region B. However, person Y sees a doctor and gets a (specific) diagnosis. Studies relying on diagnoses cannot observe person X and would therefore come to the wrong conclusion when using diagnoses as the only indicator for the true health status of the underlying population. In contrast, our data allow us to measure and model both, the health statuses of individuals X and Y as well as their usage of health care services.

Third, in addition to the individual-level demand-side measures, we exploit higher aggregated level measures that intend to model regional supply-side factors of the health care system. We observe strong regional variation in the density of physicians and hospital beds. Thus we collected administrative data on physician density as well as hospital density to model the role of the health care infrastructure. Consequently, this paper not only extensively models demand side factors on the most disaggregated level possible—the individual level—but also considers aggregated supply side factors. Finally, our approach also allows us to aggregate all representative information up to a desired regional level. Doing this, we can simulate the effect of aggregating individual-level information.

Using the microdata approach described above, we analyze geographical variation in health care utilization at the state level in Germany. For three reasons, we deliberately choose the state level as the level of analysis: First, our dataset is representative at the state level and we can rely on a sufficiently large number of observations. Second, the average German state has an area of 22,300 km^2, while the average US Hospital Referral Region (HRR) has an area of 31,500 km^2. At the same time, the average population density is higher in Germany. Third, and perhaps most importantly, the basic hospital infrastructure is financed and regulated by the 16 German federal states. Hence, we consider the German federal states as good indicators for what would be a HRR in the US.

The institutional setting in Germany has several unique features that allow us to shut down various channels that may confound the empirical results in other countries. First, there exists almost universal health care in Germany; the rate of the uninsured is less than 0.5 percent. Second, in international comparisons, the level of coverage can be considered as relatively high and the level of cost-sharing as relatively low, especially when compared to the US. Third, in international comparisons, the density of physicians and hospitals is relatively high and the hospital occupancy rate lies at around 80 percent [30]. Because of the geographic conditions, distances to providers are also shorter than in the US. Fourth, Germany has a free choice of providers and, in principle, there is no Managed Care and hence no provider networks that may additionally constrain access to care. Access barriers are thus significantly lower than in other countries. Fifth, reimbursement rates are not negotiated individually between insurers and providers, but are determined centrally, at least at a regional level. Inpatient care is reimbursed by DRGs that have regional discounts or surcharges. In other countries, all of the demand and supply-side factors just listed may vary regionally and produce regional variations in health care utilization. Lastly, we rely on data that is representative for the whole population, not just the elderly. Obviously, the behavior of the elderly or how they are treated by providers may differ from the rest of the population.

Looking at the raw, unconditional, regional health care utilization patterns, we find substantial regional variation in Germany. As for the outpatient sector, doctor visit rates vary between 90 and 120 percent of the national mean. For the inpatient sector, the number of hospital nights even varies between 65 percent and 165 percent of the national mean: Thus, inpatient sector utilization varies up to the factor 2.5 between the German states. However, not surprisingly, variation in health care utilization between individuals within the same state is much larger.

Interestingly, an inverse picture for inpatient vs. outpatient health care utilization appears: While we find above average inpatient sector utilization in the former GDR states, outpatient sector utilization is substantially below the national average in these states. However, once we control for our rich set of state-level and individual-level background information, including various measures of health and health behavior, the significant differences in health care use between the 16 German federal states disappear.

The rest of the paper is organized as follows: In Section 2 we provide a short description of our underlying dataset, the German Socio-Economic Panel Study (SOEP). Section 3 explains the underlying empirical approach and presents the results. Section 4 concludes.

2 Data

2.1 Dataset

We make use of extensive individual-level data provided by the German Socio-Economic Panel Study (SOEP). The SOEP is a representative panel study of private households. Interviews have been carried out annually since 1984. All respondents answer one main individual questionnaire covering about 150 questions on different topics such as the labor market and family situation, attitudes and perceptions as well as health. Additionally, a household questionnaire is completed by the head of the household. About 20,000 individuals from more than 10,000 households are surveyed each year. For further details, please see Wagner, Frick and Schupp (2007) [31].

The SOEP provides two main health care utilization measures: i.) the number of nights in a hospital in the last calendar year; and ii.) the number of doctor visits in the last three calendar months prior to the interview. Various subjective and (quasi-)objective individual health measures are surveyed in regular intervals. For example, the standard 5-categorical Self-Assessed Health (SAH) measure is surveyed annually. In addition, since 2002 in even numbered years, the continuous and quasi-objective SF12 measure and the objective grip strength measure have been surveyed. Furthermore, information on health-related behavior (e.g. alcohol and tobacco consumption) is available since wave W in 2006. Therefore, we restrict our analysis to the three waves: W (2006), Y (2008), and BA (2010). We use only observations without item-non-response. In total, we obtain 54,362 person-year observations from 23,167 individuals.

The SOEP routinely provides frequency and probability weights. We use these weights throughout our analysis in order to guarantee the representativeness of the data at the state level [32, 33].[3]

2.2 Outcome Variables: Health Care Utilization Measures

We exploit two dependent variables that measure outpatient and inpatient health care utilization: The SOEP asks every respondent annually: "*Were you ever admitted to a hospital for at least one night in 200X?*" Those who answer with "yes" are asked: "*How many nights altogether did you spend in the hospital last year?*" Using this information, we generate a variable *hospitalnights*. The great advantage of this measure is the representativeness, i.e., we have a general and reliable indicator of inpatient sector utilization from a representative set of respondents in every German state. Measurement errors should not dramatically bias the results; as only respondents with very long hospitalizations might not remember the exact number of hospital nights in the previous calendar year.[4]

Note that we can also exploit the health care utilization measures for the non-users, i.e., those that did not use inpatient care. Non-users provide helpful information to the econometric model since they might have a health and socioeconomic status that is comparable to individuals in other states who used inpatient care. It is important to keep in mind that the incidence of inpatient care utilization is very low and lies at around 12 percent, meaning that 88 percent of all residents do not use inpatient care. Despite the low incidence rate, Table 1 demonstrates that Germans, on average, spend 1.5 nights per year in a hospital.

[3] The usage of SOEP frequency weights guarantees representativeness at the federal level, among larger states as well as between East and West Germany. This means that state-dependent differences in participation rates—e.g., due to cultural differences—are taken care of. However, even weighted data may not be strictly representative in some smaller states such as Hamburg (1.8 m inhabitants), Bremen (0.9 m inhabitants), and the Saarland (1 m inhabitants), where the SOEP only surveys about 230, 100 and 150 people per year.

[4] As correctly pointed out by one of the referees, there is the possibility that state-dependent reporting biases in our outcome variables could confound our estimates. This is indeed true and a limitation of this study. However, while it has been shown that cross-country reporting biases—presumably due to differences in language and cultures—exist [34], we failed to find a reference proving the existence of region-dependent reporting bias within one nation. Obviously, one would need merged register and survey data to assess this issue. Such an analysis might also be a fruitful contribution to the small area variations literature.

The second dependent variable measures the utilization of the outpatient sector. The SOEP asks every respondent annually "*Have you gone to a doctor within the last three months? If yes, please state how often.*" Analogously to the inpatient health care measure, we generate a continuous variable *numberdoctorvisits*. It varies between 0 and 99, with an average of 2.5 visits per respondent (see Table 1). 70 percent of all respondents visited their doctor at least once in the last 3 months (not shown in Table 1).[5]

Clearly, the main drawback of these measures is that they are (i) general measures of health care and do not differentiate between different inpatient and outpatient types of care; and (ii) aggregated measures and do not disentangle multiple hospital stays and nights per stay. Fortunately, the SOEP also includes information on the number of stays in a given calendar year. Conditional on being hospitalized, only 25 percent of all respondents have multiple hospital stays. Since only 12 percent of all respondents are hospitalized in a given year, only 3 percent of all respondents have multiple hospital stays. In order to test whether multiple stays might confound our estimates, we run a robustness check and conduct the same analysis selecting on those with just one stay. The results are very robust and available upon request.

We also check whether regional variation in hospital mortality rates could bias our results. This could be the case since people who die after a hospital admission are obviously no longer able to participate in the survey. However, variation in hospital mortality rates pose no threat to our estimates since only about 0.3 percent of the German population die in a hospital after a hospital admission [35]. This figure translates into about 50 respondents per year—far too few to have the statistical power to substantially bias our results, even if state differences in hospital mortality rates existed.

We are able to conduct an actual additional actual robustness check that also confirms the back-of-the-envelope calculation above: On a regular basis, the SOEP Group conducts dropout studies in order to identify the status of former SOEP respondents. Using this information, we identify 623 individuals who died between 2007 and 2010.[6] 190 of these 623 individuals reported having been admitted to a hospital in calendar year 2006. In every subsequent year approximately 50 of those 190 individuals died. A separate regression reveals that the risk of dying after hospital admission does not vary significantly across federal states.[7]

Concerning outpatient care, we can neither distinguish between visits with Primary Care Physicians (PCPs) and specialists, nor can we disentangle the underlying health reason for the visit. This means that we assume that a.) the ratio between PCP and specialist visits is similar across the German states; and b.) if existent, physician-induced demand does not vary significantly across states. The latter assumption seems relatively plausible. Moreover, almost all studies in the empirical small area variation literature work under

---

[5] Note that simply multiplying this variable by the factor four would dramatically overestimate the annual number of doctor visits. The reason is that about two thirds of all interviews are carried out in the first quarter of a year and there is a clear seasonal pattern in doctor visits. However, the crucial point is that the focus on the last three months substantially reduces measurement errors. Moreover, we control for months fixed effects in our regression models.

[6] The SOEP-LIFESPELL file contains the dropout information.

[7] Only the Berlin state dummy is significant at the 5 percent level implying that the risk of death after a hospital admission is significantly lower in Berlin than in the reference state, Schleswig Holstein. All other state dummies are not significant at the 10 percent level.

this assumption. Note that geographic variation in physician-induced demand might either increase or decrease "natural" geographic variation in health care utilization, depending on how healthy and unhealthy people sort across regions with different degrees of physician-induced demand. The former assumption— that the ratio of PCP vs. specialist visits is evenly distributed across regions— might be more critical. 1994 was the only post-reunification year when the SOEP surveyed the number of PCP *and* specialist visits. We exploit that information and created an outcome variable that measures the ratio of specialist vs. PCP visits. Running a robustness check of that ratio on a set of state dummies did not yield any empirical evidence that the *ratio* of specialist vs. PCP visits differed significantly between the German federal states in 1994.[8]

2.3 Individual and State-Level Control Variables

By taking into account demographic factors, educational characteristics, the labor market status, the health status and health-related behavior as well as state-level indicators of the health care infrastructure, we net out all differences in health care utilization on the state level that can be traced back to these factors.[9] Table 1 shows the summary statistics of all variables used.

Individual-level socio-economic measures

The ***demographic factors*** that we use are age, age squared, a male gender dummy, a dummy for being married, and the number of children under 14 in the household. Table 1 shows that average age is about 50 years and that about half of our sample is male. The majority of respondents are married and, on average, 0.36 children live in SOEP households.

In terms of ***education and labor market participation,*** we control for whether an individual completed vocational training or holds a university degree, the work status of the respondent, and the monthly equivalent household income.[10] About a fifth holds a university degree (*master level or higher*) and has no college degree, respectively. The remaining 60 percent completed vocational training, which is equivalent to a college degree in the US. Slightly less than half of our sample works and the average

[8] Only one of the state dummies is significant at the 10 percent level and yields that the specialist-PCP ratio in Berlin is significantly higher than in the reference state, Schleswig Holstein.

[9] Please note that, strictly speaking, all right-hand side time-variant information is measured at the time of the interview, while all left-hand side information is surveyed retrospectively. This is unlikely to be an issue for several reasons: First, the time span is relatively short and the nature of the time-variant information surveyed is relatively stable across time. Second, in case of the respondents' health status, difference in the time structure might be an issue since health can be seen as a time-variant outcome of health care utilization. However, please keep in mind that we only intend to net out correlations and do not intend to provide evidence for causality. Consequently, even if health on the right-hand side of the equation was a time-varying outcome variable, the framework would still make sense. In that case, it could be interpreted as measuring whether different levels of health care utilization produce different health outcomes. If we found significant state-level effects, it could mean that the effectiveness of treatments differ significantly across states. Finally, we run a robustness check and assign health care utilization in t+1 to the time variant information provided in t, in case that the respondent was interviewed in t and t+1. The results remain stable and are available upon request.

[10] The monthly equivalence household income uses the OECD-modified scale and assigns households a value of 1 to the first adult, 0.5 to each additional adult and 0.3 to each additional child. Further details are provided in *OECD Project on Income Distribution and Poverty (2009)* [36].

equivalent gross household income per capita is about €1,800 per month. Approximately 15 percent in our sample are privately insured.[11]

Individual-level health and health behavior measures

We make use of two different types of individual ***health measures***. They are presumably the most important individual-level predictors of differences in health care utilization. First, we use the standard (subjective) 5-categorical Self-Assessed Health (SAH) measure and generate two different dummy variables. *SAH good health* indicates the share of respondents (9 percent) who rate their own health as excellent. *SAH bad health* includes the 18 percent who state that their health status would be fair or poor. While SAH is a rough and subjective measure, it is shown to be a very good predictor of the true health status [37]. Moreover, we control for all factors that are shown to vary systematically with the self-assessment of respondents' health, i.e., age, gender, labor market status, and income [38].

Second, we make use of the generic health measure SF12. The continuous SF12 weights and aggregates the answers to twelve health questions into a physical health (*pcs*) and a mental health (*mcs*) summary scale using a specific algorithm. Developed by public health scientists, the SF12 belongs to the group of "health-related quality of life measures." Although SAH is part of the SF12, research shows that SF12 is a more objective health measure than the raw SAH, with the advantage that it "can be used to compare the health of different groups, for example, the young and the old or the sick and the well [39]." We call the continuous SF12 health measures "quasi-objective" [40]. In the standard SOEP version*, pcs* and mcs take on continuous values between 0 and 100, have mean 50, and a standard deviation of 10 (see Table 1). A detailed description of the algorithm and an overview over the differences between the original "SF12v2™ Health Survey" and the SOEP version can be found in Andersen et al. (2007) [41].

Concerning ***health-related behavior***, we control for alcohol and tobacco consumption, the Body Mass Index (BMI), and regular exercising. Alcohol consumption is captured through dummy variables. In the SOEP questionnaire, the participants are asked to state how often they drink wine, sparkling wine, beer, spirits and mixed drinks. If an individual states that they do not drink any alcohol at all, *no alcohol consumption* is assigned the value "1", otherwise it is "0". In contrast, if an individual states to drink any kind of alcohol on a regular basis, *regular alcohol consumption* takes on the value "1", otherwise "0" [42]. 17 percent drink alcohol on a regular basis and 13 percent never drink any alcohol. The dummy variable *Smoker* captures whether the respondent consumes tobacco (cigarettes, cigars or pipes). It takes on the value "1" for about 30 percent of the respondents in our sample. The *BMI* is measured as body weight in kilogram divided by the squared height in meters. The average BMI lies slightly below 26. *Regular sports* is a

[11] Germany has a two-tier health care system. The great majority of the population is insured under the public scheme. Only higher income earners, self-employed people, and civil servants have the right to opt out of the public scheme and insure their health risks privately. Outpatient reimbursement rates in the private system are higher than in the public system, but are also centrally determined.

dummy variable that takes on the value “1” for the 37 percent of the sample who report doing sports or exercise at least once a week.[12]

[Insert Table 1 about here]

State-level measures of the health care infrastructure

In order to explain systematic state-level differences in the health care infrastructure, we incorporate two different measures into our econometric analysis. The data is provided by the FEDERAL INSTITUTE FOR RESEARCH ON BUILDING, URBAN AFFAIRS AND SPATIAL DEVELOPMENT (2012) [43].

Panel B of Table 1 shows that the first measure is an indicator of the outpatient health care infrastructure, *Inhabitants per physician*. This supply-side measure varies by up to a factor of two across German states. On average, Germany has 625 residents per outpatient care physician—this is a high number in an international comparison, e.g., significantly higher than in the US [44].

The second measure is an indicator of the inpatient health care infrastructure and is called *Hospital beds per 10,000 inhabitants*. Germany has on average 60 hospital beds per 10,000 inhabitants. Here, the variation is even larger as compared to the physician density and varies up to a factor of six across states. And again, the density of hospital beds is very high in an international comparison. Among OECD countries only Japan has a higher rate [45].

3 Empirical Approach and Results

3.1 Illustrating Raw State-Level Variation in Health Care Utilization

In a first step, we simply summarize and illustrate unconditional levels of health care utilization, state by state. All measures are weighted by SOEP provided frequency weights and are representative.

Inpatient care

Figure 1 plots the average number of hospital nights in the 16 German federal states. As the legend shows, it ranges from 1.05 through 2.73 nights per person, and hence varies up to 2.7 times between states. The SAV literature from the US observes that some regions consume up to twice as much health care than other regions, without showing better health outcomes. Figure 1 is the German equivalent to this finding. And, as in the US, the average health status in the high consumption states is lower, not higher, when compared with low consumption states (results not shown). However, these state-level differences in health are not statistically significant from zero.

[12] Information on the frequency of sports and exercise is not available for 2006 and 2010. However, it is plausible that this habit changes rather slowly. We assume that it is constant during the time period in our analysis and impute the observations for 2006 and 2010 with the observed value in 2008.

[Insert Figure 1 about here]

When looking at Figure 1, we notice one eye-catching pattern: The five former GDR states in the North-East of Germany (black solid line: *Mecklenburg-Western Pomerania, (East) Berlin, Brandenburg, Saxony, Saxony-Anhalt, Thuringia*) show systematically higher levels of inpatient care. This could be due to several reasons. One obvious and plausible reason is that four decades of life under communism with substantially higher levels of (air) pollution, poorer diet as well as hygiene conditions have had long-lasting health effects that manifest themselves through higher hospitalization rates. Another reason could simply be migration. Since the German reunification in 1990, a total of 3 million, primarily young and healthy, East Germans have migrated to West Germany.

[Insert Figure 2 about here]

Figure 2 includes the same information as Figure 1, but illustrates it in a different manner. It shows the average inpatient health care utilization for the 16 German states, relative to the national mean. As can easily be seen, there is substantial variation in hospitalization rates, ranging from 65 to 165 percent of the national mean. Again, we observe that five of the six former East German states are substantially above average inpatient care utilization by their residents.[13] In a strict statistical sense, and when using Schleswig-Holstein[14] as a reference state, Brandenburg's and Thuringia's inpatient care levels are significantly higher at the 5 percent level. Saxony-Anhalt's and Mecklenburg-Western Pomerania's levels are also significantly higher at conventional statistical levels, i.e., at the 10 percent level. Berlin's hospitalization rate is only significantly higher at the 16 percent level, which is, however, very likely to be due to the lower number of SOEP respondents in Berlin and, thus, lower statistical power.[15]

Outpatient Care

Figure 3 again shows the German map, but now plots outpatient care utilization rates. Here the state-level variation is less pronounced, but still ranges from 2.19 up to 3.01 doctor visits over the last three months before the interview. One striking result and a key finding of this paper is that we find the mirror picture to hospitalizations: While hospitalization rates in the former GDR states were significantly and systematically higher than in the rest of Germany, the opposite is true for doctor visits. In case of outpatient care, the former GDR states show lower utilization rates. One exception is Berlin, which is a metropolitan area and has the highest physician density of all German states (413 residents per physician), as well as the highest population density.

[Insert Figure 3 about here]

[13] Saxony is the excepton.
[14] The German state in the very north of Germany.
[15] In addition, note that we do not distinguish between East and West Berlin. If it was true that hospitalization rates in East Berlin are significantly higher than in West Berlin, we would not necessarily be able to detect it in this analysis.

The fact that the former GDR states display higher hospitalization rates, but lower outpatient visit rates than the rest of Germany could be an artifact of various underlying mechanisms. One appealing and plausible explanation, which would be highly relevant for health policymakers, is the following: To some degree, outpatient and inpatient care are substitutes. At the same time, regular and continuous primary care physician visits serve as a form of preventive care and may prevent severe medical conditions that would otherwise manifest itself in the long-run and require a hospital stay. These conditions could be avoidable through timely and appropriate outpatient care. The fact that we observe below average outpatient utilization and above average inpatient utilization fits that explanation perfectly. In addition to the preventive argument, outpatient treatments are less costly than inpatient treatments.

If it is true that outpatient care prevents inpatient care, the policy conclusion of this empirical finding would be to identify the reasons for the low outpatient utilization patterns in East Germany and implement policy measures accordingly. The former East German states—with the exception of Berlin—have an outpatient physician density that is below average. Policy-makers have long debated how to fight the shortage of Primary Care Physicians (PCPs) in rural areas; most frequently in East Germany. Effective January 2012, a new health care reform was implemented designed to target the shortage of PCPs in rural German areas.[16] The main measures to achieve this goal are monetary incentives for PCPs who are willing to practice in regions that are identified as "shortage regions." However, the reform is criticized as it does not target the oversupply of physicians in some (metropolitan) areas.

Although it is likely that the low outpatient care utilization in the former GDR states is mainly due to a low physician density, a simple correlation does not prove causality. Theoretically, other GDR-related factors could be the driving forces of this correlation. Another explanation could be the high unemployment rate and low incomes in these states. Although copayments for PCPs are moderate, hardship clauses exist, and research finds mixed evidence concerning the effectiveness of German outpatient copayments [45-47] as people in poor economic conditions might postpone doctor visits because of copayments.

Interestingly, if we just plot the incidence of doctor visits, i.e., whether people saw a doctor at all, the clear East-West pattern vanishes (Figure not shown). This is evidence in line with an explanation that emphasizes physician density as a driving force of the low utilization rates. In the econometric models below, we incorporate physician density as one explanatory factor.

[Insert Figure 4 about here]

Figure 4 is the analogue to Figure 2 above: state-level doctor visits are expressed relative to the national mean. It is easy to see that the five former GDR states, excluding Berlin, show an outpatient utilization rate that lies between 87.2 (Saxony) and 94.4 percent of the national mean rates. Two of these differences

[16] This law is the "*Gesetz zur Verbesserung der Versorgungsstrukturen in der Gesetzlichen Krankenversicherung (GKV-Versorgungsstrukturgesetz, GKV-VStG*)."

are statistically significant at conventional statistical levels (1.6 percent for Saxony and 8.8 percent for Saxony-Anhalt) and the other two differences are statistically significant at the 13 and 17 percent levels.

Also note that, not surprisingly, the individual-level variation in health care utilization is substantially larger than the variation at the state level. As Table 1 illustrates, in case of outpatient care, it varies up to the factor 99 and in case of inpatient care, up to the factor 280.[17]

3.2 Analyzing State-Level Variation in a Regression Framework

Econometric model

In the following, we incorporate our rich set of individual-level and state-level explanatory variables into an econometric model in order to adjust the state-level differences to these factors (see Section 2.3). We run the following model by means of an OLS regression:

$$\begin{aligned} y_{it} = \beta_0 + \beta_1 Demographics_{it} + \beta_2 Education_{it} + \beta_3 LaborMarket_{it} + \beta_4 Health_{it} + \beta_5 HealthBehavior_{it} \\ + \beta_6 \Pr oviderDensity_{st} + \beta_7 state_s + \beta_8 month_m + \beta_9 year_t + \varepsilon_{it} \end{aligned}$$

where y_{it} stands for either the number of hospital nights or the number of doctor visits (see Section 2.2), $Demographics_{it}$ is a vector of individual-level demographic control variables, $Education_{it}$ is a vector of individual-level educational control variables and $LaborMarket_{it}$ is a vector of labor market covariates (see Section 2.3 for further details). $Health_{it}$ and $HealthBehavior_{it}$ are potentially very important driving forces of differences in health care utilization. In our framework, we include two subjective and two quasi-objective health measures. Moreover, we consider health behavior measures for alcohol and tobacco consumption, obesity as well as sports participation. $\Pr oviderDensity_{st}$ includes the hospital bed density as well as the physician density and takes account of the supply side, i.e., the health care infrastructure. Both variables vary across time and between states. $state_s$ is a set of state fixed effects, $month_m$ a set of months fixed effects and $year_t$ includes two year dummies. ε_{it} is the error term, as usual.

Note that the state, month and time-level controls are dummy variables, not random components, i.e., we include 15 state dummy variables, 11 month dummy variables and 2 year dummy variables. The only random component in this model is the error term. Without further controls, the coefficients of our 15 state dummy variables would simply display deviations in health care use from the suppressed references state, i.e., we obtain the regression equivalent to figures 1 to 4. The model could also be formulated in a multilevel (or hierarchical) framework where individuals are nested within states. In this

[17] Taking into account that people who are seriously sick cannot participate in the SOEP, the true variation in case of hospitalizations ranges up to the factor 365. The same would be true for outpatient care under the assumption that some people have to visit their doctor every day.

case, the state effects are modeled as random deviations from the overall constant, i.e. random effects. Such a multilevel model introduces state-specific variation in addition to the error term.

Also note that neither this study nor the econometric model claims or intends to provide "causal effects." It solely deals with correlations, as does the large majority of the SAV literature. As noted, we start off with a parsimonious model that only includes state dummies. Then we further add time dummies and, lastly, we net out all individual-level variation in utilization that is systematically correlated with a rich array of individual-level observables. Clearly, there exist unobservables, *z*, that may be correlated with both utilization, *y,* and our individual-level covariates, *x*. This leads to omitted variable bias, which does not allow us to interpret the coefficients of the individual-level covariates, *x*, as causal effects. The impact of the unobservable *z* on *y* is captured by the *x* coefficient. This is, however, no threat to our research design since we are not primarily interested in the *x* coefficients but in netting out as much of the individual-level variation in *y* as possible.

The other main endogeneity issue is clearly reverse causality, e.g., the fact that *y*— in this case health care utilization—may impact *x*—for example health. Since we are primarily interested in the existence of a significant correlation between health care utilization and states, conditional or unconditional, reverse causality is neither a threat to our research design nor to the validity of our conclusions (see footnote 7 for more details).

Conditional state-level differences in health care utilization

We now include the sets of individual-level, state-level, and time-level covariates and run the model as outlined above.[18] The results are displayed in Table 2 below. Each column represents one model. The detailed coefficient estimates for the state-level dummy variables are suppressed to save space.[19] The findings can be summarized as follows:

When netting out state level differences that go back to the covariates described above and in Section 2.3, we cannot reject the null hypothesis that there are no state level differences in health care utilization—at least not in strict statistical terms.

In the final model for hospitalization rates (Model 1c), we only find two significant state level dummies (*Hesse, Hamburg*). In the final model for doctor visits (Model 2c), only the dummies for *Rhineland-Palatinate* and *Lower Saxony* are significant at the 10 percent level. In addition, depending on the model specification and the sets of covariates that we include in the models, there is not even consistency in the state dummies that turn out to be significant. Moreover, given that we incorporate 15 state dummies and apply a statistical error probability of 10 percent, it is quite natural to obtain 1-2 significant dummy variables. All

[18] The standard errors are adjusted to account for heteroscedasticity at the individual level.
[19] The detailed results are available upon request from the authors.

in all, we conclude that we do not observe any significant state-level differences in health care utilization once we control for a rich set of individual-level, state-level, and time variables.[20]

We also run multilevel models that model the state effects as random deviations from the overall constant (results not shown). The findings from these alternative models provide further evidence for the conclusion that there are no systematic and economically significant state-level differences in health care use, once a rich set of individual-level background characteristics of the residents are controlled for. Multilevel models can be summarized by the intraclass correlation coefficient, which is a ratio of state-specific variation to total random variation (i.e. the sum of state-specific variation and residual variation). This correlation coefficient can be regarded as a measure of similarity for individuals within the same state. In the model for hospitalization rates this coefficient is approximately 0.01 and in the model for doctor visits this coefficient is 0.05. Both numbers are extremely small, indicating that there is little to no similarity between individuals within the same state that would result in systematic group-specific effects. In other words, there are no significant state-level effects.[21]

Using micro level data also allows us to conduct another exercise: We aggregate all variables up to the state-level and run the same regression models with just one observation per state and year, i.e., 48 observations in total. As such, we simulate the effect of aggregating data at a higher level. Our state-level results are robust to this exercise and almost all of them are insignificant.

Factors that account for the observed differences in unconditional health care utilization rates

Let us now turn to the factors that are responsible for the observed differences in the raw health care utilization rates. Table 2 displays our regression models and the coefficient estimates for the individual-level and state-varying control variables. Again, our model does not allow for causal interpretation of the coefficient estimates.

The first three columns show the results for hospitalization rates and the last three columns show the results for doctor visits. The columns—and hence models—only differ by the inclusion of different sets of covariates.

[Insert Table 2 about here]

We start with demographics. Men stay on average half a day longer in the hospital when compared to women. This is a huge effect if we consider that only 12 percent of the population has hospital stays in a given year and that the mean annual number of hospital days is 1.5. Married people spend significantly

[20] This conclusion is robust with respect to different model specifications. For example, we run models 1c and 2c as negative binomial regressions to account for the fact that our dependent variables can be interpreted as count data. Also, since the distribution of the dependent variables is highly skewed with a long right tail and overdispersion (variance exceeds the mean), count data models might fit the distributional properties better than OLS models [48]. However, the results and main conclusions are robust to alternative count data specifications. These are available upon request.

[21] In this multilevel model, group-specific effects are estimated as deviations from the overall constant, i.e. they have a mean of zero. These deviations are summarized through their standard deviation, which in turn is used for the calculation of the intraclass correlation coefficient. Thus a small standard deviation implies that the group-specific effects are very close to zero.

fewer days in hospitals than the unmarried. What is interesting to observe is that the signs of these associations completely switch for the outpatient sector (Models 2a to 2c): Here we find that men visit outpatient physicians less often than women, but married people visit them more often than the unmarried. This resembles the mirror picture for outpatient and inpatient care that we observed in the case of the East-West differential. It supports the hypothesis that the two types of care are to some degree substitutes and that more outpatient care may prevent hospitalizations later in life.

There is also some correlation of the utilization rates with labor market status and educational level. For example, non-working people are more often hospitalized. Interestingly, income is negatively correlated with the number of hospital nights (Model 1a), but once we control for health (Model 1c), the sign of the correlation switches and we find a positive correlation. Note that the latter is true for both inpatient and outpatient care. This finding might be interpreted as evidence for additional consumption of expensive, but medically not necessary, out-of-pocket medical care. Another interesting and plausible finding is that privately insured people visit their doctors more often, but are not hospitalized as often. Outpatient reimbursement rates for the privately insured are higher than for the publicly insured in Germany.

Now we assess the relative importance of the different sets of covariates by comparing the R^2 of our different models. The R^2 is a statistical measure that tells us how much of the variation of the outcome variable is explained by the covariates of the econometric model. In micro data models, the R^2 is typically fairly low. A model that explains 20 percent of the variance of the outcome variable can be regarded as a rich, well-fitted model. If we run a plain regression model that solely incorporates state dummies, i.e., the regression equivalent to Figures 1 to 4, the R^2 is 0.001 for the inpatient and 0.002 for the outpatient sector. This means that less than 0.2 percent of variation in health care utilization can be explained by simple state-level dummies. Additionally controlling for monthly and yearly time shocks as well as the physician and hospital density doubles the explanatory power of the models. However, it remains very low—below 0.5 percent. These model diagnostics show that solely relying on state-level factors and time trends to explain health care utilization of individuals yield empirical models with a poor fit and insufficient explanatory power, even though some coefficient may be significant in statistical terms.

As seen in Table 2, adding individual-level control variables for demographic, educational, and labor market characteristics substantially improves the model fit. The R^2 for both models increases by a factor of 20 and the models explain now about 2.5 and 5 percent of the individual-level variation in health care utilization, respectively. Note that adding further socio-economic explanatory variables further increases the model fit. The R^2 in our preferred model specifications 1c and 2c—we incorporate all covariates described in Section 2 in these models— amounts to 6.5 percent for the inpatient and 19 percent for the outpatient sector. As already discussed, although large parts of the variance in the final models remain unexplained, these models do not yield any evidence that general state-level effects over and above the other control indicators exist and are systematically and significantly associated with health care utilization.

Let us now turn to presumably the most important driving forces of health care utilization patterns: health. Looking at models 1b and 2b in Table 2, there is no doubt that the individual health status is, by far, the most important correlate of health care utilization. For the two subjective and the objective general health measures, we observe strong and highly significant correlations in all our models. This is also reflected in the R^2. General individual-level measures for health substantially improve the model fit and do a great job of explaining individual-level differences in health care utilization. This is not surprising but underscores the importance of controlling for health in such analyses. The controls in models 1b and 2b account for over 90 percent of the explanatory power of our empirical models as indicated by the R^2.

Measures of individual health behavior also constitute an important set of control variables, although selection issues become apparent. For example, the fact that *no alcohol consumption* is strong and positively correlated with both, inpatient and outpatient care utilization, is probably an artifact of the observation that seriously sick people are more likely to not drink any alcohol at all. Likewise, such a selection story explains the negative correlation between *regular alcohol consumption* and doctor visits (Model 2c) as well as between smoking status and health care utilization. Alternatively, one could argue that heavy drinkers do not care much about their health, which is why they are less likely to visit doctors.

Lastly, people who exercise regularly visit their doctors more often. There are several possible reasons, including i.) exercising causes more body injuries; ii.) these people care more about their bodies; or iii.) their doctors recommended them to exercise more.

Looking at Model 1c and the state-level controls that represent the health care infrastructure, we find that physician density is negatively correlated with hospitalization rates. Since the physician density measure only varies between the 16 states and (slightly) across years, we lack statistical power, but the correlation is still significant at the 13 percent level if we incorporate only the infrastructure indicators, state, year, and month fixed effects as controls (not shown). However, the size of the correlation is relatively strong: one additional inhabitant per physician increases average hospital nights by 0.013. This would be equivalent to 1 million additional hospital nights for the whole of Germany. And again, this negative correlation between physician density and hospital nights is in line with the presumption that timely outpatient visits may prevent later inpatient hospital stays to some degree. The correlation between the physician density and the number of doctor visits is weak and not statistically significant (Model 2c). However, if we aggregate all variables up to the federal state level and rerun Model 2c, we find that the physician density and the number of outpatient visits in a state are strongly and negatively correlated. The coefficient estimate is -0.00615 and significant at the 3 percent level. This would translate into up to 500,000 fewer outpatient care contacts for every 200 missing outpatient care physicians or up to 2,200 contacts per physician and year.[22] This empirical evidence supports the story of physician-induced

[22] Note that we overestimate the relationship since two-thirds of all interviews are carried out during the first quarter of a year and doctor visits are highly cyclical with respect to the season.

demand. However, as noted, this simple correlation does not prove the existence of physician-induced demand.

When it comes to the hospital bed density and utilization patterns, we do not find any evidence that the inpatient care infrastructure is correlated with the number of doctor visits. In contrast, a high hospital bed density is negatively and significantly correlated with hospitalizations. This means that we observe, on average, fewer hospitalizations in states with more hospitals.

Conclusion

This paper makes use of a microdata approach to analyze variation in regional health care utilization patterns in Germany. The main advantage of this approach is the representativeness of the data, the consideration of non-users, and the richness of the socio-economic background information. The latter can be incorporated into econometric models in order to consider a vast array of demand and supply-side factors which might explain regional differences in health care utilization.

The unconditional and raw utilization patterns reveal some interesting findings for Germany. First, there is substantial regional variation in utilization patterns across the 16 German federal states. Second, inpatient sector utilization varies more across states than outpatient sector utilization. Hospitalization rates vary from 65 to 165 percent of the national mean, i.e., up to a factor of 2.7. In contrast, doctor visit rates only vary between 90 and 120 percent of the national mean. Third, inpatient and outpatient care utilization mirrors each other. While hospitalization rates for the former GDR states lie systematically and significantly above the national average, doctor visit rates lie systematically and significantly below the national average. This provides, third, empirical evidence for the presumption that a timely and regular use of outpatient care may be effective in avoiding the onset of serious diseases that require hospital stays.

We then adjust the state-level differences in health care utilization that can be traced back to individual-level state population differences in demographics, employment and educational characteristics as well as health and health behavior. Moreover, we adjust for state-level differences in the health care infrastructure, months, and year effects. If we do that within a regression framework, all regional differences in health care utilization patterns vanish and we no longer observe significant differences.

4 References

[1] Wennberg, J. E., Gittelsohn, A. (1973): "Small area variations in health care delivery," Science, 182: 1102-1108.

[2] Fisher, E. S., Wennberg, D. E., Stukel, T. A., Gottlieb, D. J. (2004): "Variations in the longitudinal efficiency of academic medical centers," Health Affairs, published online before print, October 2004, doi: 10.1377/hlthaff.var.19, last accessed on June 16, 2012.

[3] Chandra, A., Staiger, D. O. (2007): "Productivity Spillovers in Health Care: Evidence from the Treatment of Heart Attacks," Journal of Political Economy, 115: 103-140.

[4] Baicker, A., Chandra, A. (2004): "Medicare Spending, The Physician Workforce, And Beneficiaries' Quality Of Care," Health Affairs, published ahead of print April 7, 2004, doi:10.1377/hlthaff.w4.184, last accessed on June 13, 2012.

[5] Baicker, A., Chandra, A., Skinner, J. S., Wennberg, J. E. (2004): Who You Are And Where You Live: How Race And Geography Affect The Treatment Of Medicare Beneficiaries, Health Affairs, published ahead of print October 7, 2004, doi:10.1377/hlthaff.var.33, last accessed on June 13, 2012.

[6] Wennberg, J. E., Fisher, E., Skinner, J. S. (2002): "Geography and the debate over Medicare reform," Health Affairs, 21 (Suppl. Web Exclusive): W96–114.

[7] Wennberg, J. E., Barnes, B. A., Zubkoff, M. (1982): "Professional uncertainty and the problem of supplier-induced demand," Social Science and Medicine, 16: 811-824.

[8] Wennberg, J. E. (2004) "Practice variations and health care reform: connecting the dots," Health Affairs, Web Exclusive, published ahead of print October 7, 2004, doi:10.1377/hlthaff.var.140.

[9] Jürges, H. (2007):"Health Insurance Status and Physician-Induced Demand for Medical Services in Germany: New Evidence from Combined District and Individual Level Data," SOEPpapers 8, DIW Berlin.

[10] Zweifel, P., Breyer, F., Kifmann, M. (2009): *Health Economics*, second edition, Springer.

[11] Folland, S., Stano, M. (1989): "Sources of small area variations in the use of medical care," Journal of Health Economics, 8: 85-107.

[12] Stano, M. (1993): "Evaluating the policy role of the small area variations and physician practice style hypotheses," Health Policy, 24: 9-17.

[13] Molitor, D. (2011): "The evolution of physician practice styles: evidence from cardiologist migration, mimeo, http://economics.mit.edu/files/7301 last accessed on February 7, 2013.

[14] Epstein, A. J., Nicholson, S. (2009): "The formation and evolution of physician treatment styles: an application to cesarean sections," Journal of Health Economics, 28: 1126-1140.

[15] Moore, F. D. (1985): "Small area variations studies: illuminating or misleading?" Health Affairs, 4: 96-101.

[16] Riva, M.; Gauvin, L., Barnett, T. A. (2007): "Toward the next generation of research into small area effects on health: a synthesis of multilevel investigations published since July 1998." Journal of Epidemiology & Community Health, 61: 853-861.

[17] Dartmouth Atlas of Health Care (2012), available at www.dartmouthatlas.org, last accessed on May 2, 2012.

[18] Westert, G. P., Groenewegen, P. P. (1999): "Regional disparities in health care supply in eleven European countries: does politics matter?" Health Policy, 47: 169-182.

[19] Filippini, M., Masiero, G., Moschetti, K. (2006): "Socioeconomic determinants of regional differences in outpatient antibiotic consumption: Evidence from Switzerland," Health Policy, 78: 77-92.

[20] Costa-Font, J., Pons-Novell, J. (2007): "Public health expenditures and spatial interactions in a decentralized national health system." Health Economics, 16: 291–306.

[21] Knickman, J. R. (2004): "The Dartmouth data: moving from analysis to action," Health Affairs, published ahead of print October 7, 2004, doi:10.1377/hlthaff.var.121, last accessed on June 13, 2012.

[22] Mercuri, M., Natarajan, M. K. M, Norman, G., Gafni, A. (2012): "An even smaller area variation: Differing practice patterns among interventional cardiologists within a single high volume tertiary cardiac centre," Health Policy, 104: 179-185.

[23] Dunn, A. C., Shapiro, A., Liebman., E. (2012):"Geographic Variation in Commercial Medical Care Expenditures: A Decomposition Between Price and Utilization," BEA Working Papers 0075.

[24] Treurniet, H. F., Boshuizen, H. C., Harteloh, P. P. M. (2004): "Avoidable mortality in Europe (1980-1997): a comparison of trends." Journal of Epidemiology & Community Health, 58: 290-295.

[25] Hirst, N., Gordon, L., Gies, P., Green, A. C. (2009): "Estimation of avoidable skin cancers and cost-savings to government associated with regulation of the solarium industry in Australia," Health Policy, 89: 303-311.

[26] Sundmacher, L., Busse, R. (2011): "The impact of physician supply on avoidable cancer deaths in Germany. A spatial analysis," Health Policy, 103: 53-62.

[27] Roll, K. (2012): "The influence of regional health care structures on delay in diagnosis of rare diseases: The case of Marfan Syndrome," Health Policy, 105: 119-127.

[28] Sundmacher, L., Gaskins, M., Hoffman, K, and Busse R. (2012). "Spatial distribution of avoidable cancer mortality." Journal of Public Health, 20: 279-288.

[29] Stargardt, T., Schreyögg, J. (2012): "A framework to evaluate the effects of small area variations in healthcare infrastructure on diagnostics and patient outcomes of rare diseases based on administrative data," Health Policy, 105: 110-118.

[30] Gesundheitsberichterstattung des Bundes (2012), www.gbe-bund.de, last accessed on June 20, 2012.

[31] Wagner, G. G., Frick, J. R., and Schupp, J. (2007): "The German Socio-Economic Panel Study (SOEP) - evolution, scope and enhancements." Journal of Applied Social Science Studies (Schmollers Jahrbuch), 127: 139-169.

[32] Spieß, M.; Kroh, M.; Pischner, R.; Wagner, G. G. (2008): "On the Treatment of Non-Original Sample Members in the German Household Panel Study (SOEP): Tracing, Weighting, and Frequencies," SOEPpapers 98, DIW Berlin.

[33] Kroh, M. (2011): "Documentation of Sample Sizes and Panel Attrition in the German Socio Economic Panel (SOEP) (1984 until 2010)," SOEP Data Documentation 59, DIW Berlin.

[34] Jürges, H. (2007). True health vs. response styles: Exploring cross-country differences in self-reported health. *Health Economics*, 16: 163–178.

[35] Federal Statistical Office (2012): Krankenhausstatistik Teil II: Diagnosen der Krankenhaushauspatienten. http://www.forschungsdatenzentrum.de/bestand/krankenhaus/index.asp

[36] OECD Project on Income Distribution and Poverty (2009): *What are equivalence scales?* Availabe online at http://www.oecd.org/dataoecd/61/52/35411111.pdf, last accessed on April 30, 2012.

[37] McGee, D. L., Liao, Y., Cao, G., Cooper, R. S. (1999): "Self-reported health status and mortality in a multiethnic US cohort." American Journal of Epidemiology. 149, 41–46.

[38] Ziebarth, N. R. (2010): "Measurement of Health, Health Inequality, and Reporting Heterogeneity," Social Science & Medicine, 71: 116–124.

[39] RAND (1995): *User's Manual for the Medical Outcomes Study (MOS) Core Measures of Health-Related Quality of Life.* (by Hays, R. D., Sherbourne, C. D, Mazel, R.). RAND Monograph Report, Santa Monica,

RAND Corporation. Available at: http://www.rand.org/pubs/monograph_reports/MR162.html, last accessed on February 2, 2012.

[40] Frick, J. R., Ziebarth, N.R. (2012): "Welfare-Related Health Inequality: Does the Choice of Measure Matter?" The European Journal of Health Economics, forthcoming.

[41] Andersen, H. H., Mühlbacher, A., Nübling, M., Schupp, J., and Wagner, G. G. (2007): "Computation of standard values for physical and mental health scale scores using the SOEP version of sf12v2." Journal of Applied Social Science Studies (Schmollers Jahrbuch), 127, 171-182.

[42] Ziebarth, N. R., M. M Grabka (2009): "In Vino Pecunia? The Association Between Beverage-Specific Drinking Behavior and Wages," Journal of Labor Research, 30, 219-244.

[43] Federal Institute for Research on Building, Urban Affairs and Spatial Development (2012): *INKAR 2011 Indicators, Maps and Graphics on Spatial and Urban Monitoring*, CD-Rom.

[44] OECD Health Data (2011): http://stats.oecd.org, last accessed on June 25, 2012.

[45] Augurzky, B., Bauer, T., Schaffner, S. (2006):"Copayments in the German Health System – Do They Work?" RWI Discussion Papers 0043.

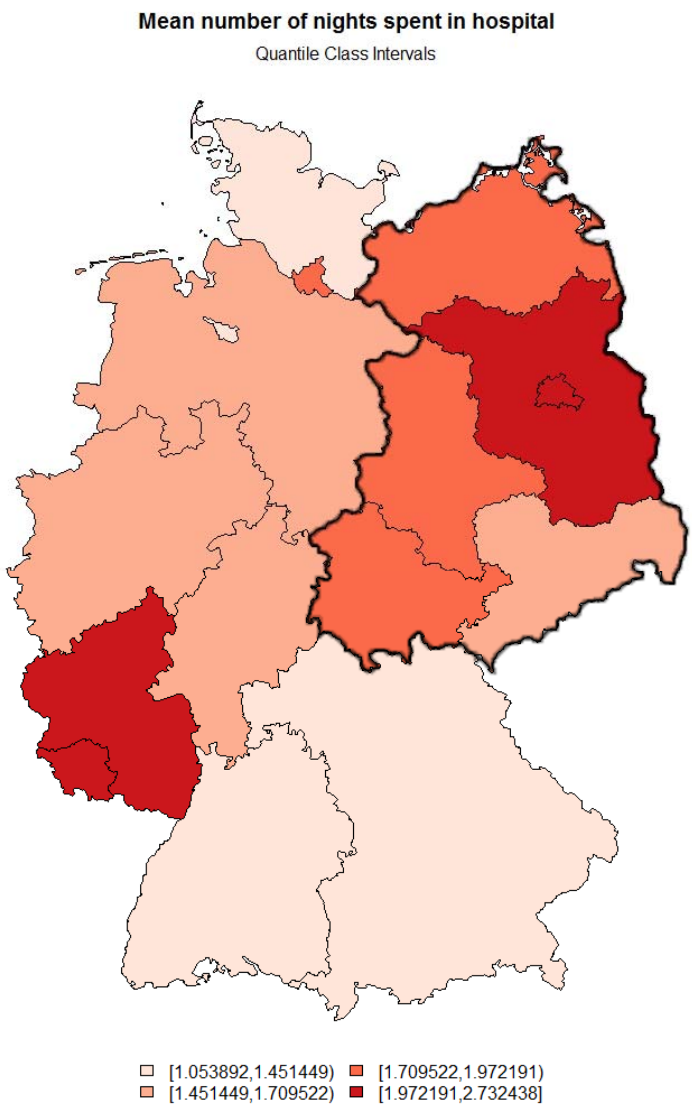
[46] Farbmacher, H. (2009): "Copayments for doctor visits in Germany and the probability of visiting a physician - Evidence from a natural experiment," Discussion Papers in Economics 10951, University of Munich, Department of Economics.

[47] Schreyögg, J. and Grabka, M. M. (2010): "Copayments for ambulatory care in Germany: a natural experiment using a difference-in-difference approach," The European Journal of Health Economics, 11: 331-341.

[48] Winkelmann, R. (2008): *Econometric Analysis of Count Data*, 5th edition, Springer.

Tables and Figures

Figure 1: Annual Number of Hospital Nights by German States (Frequency Weighted)

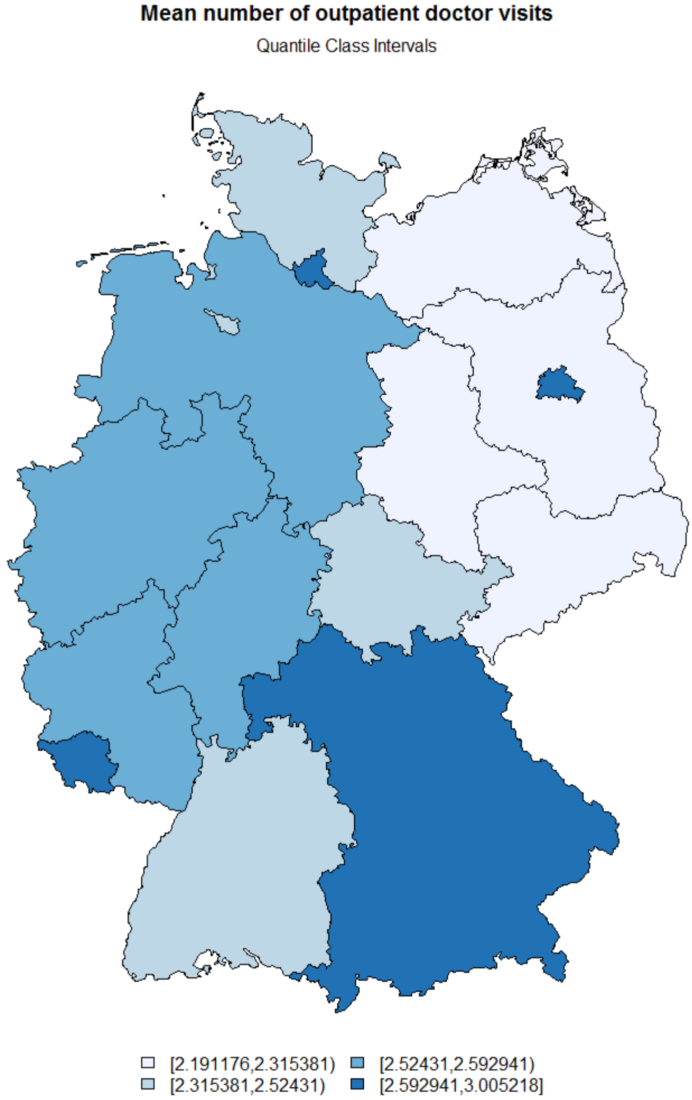


Sources: SOEP v27, own calculations. The mean number of nights spent in a hospital during the last year is displayed. The solid black line highlights the territory of the former GDR. The values of the variable are divided into four classes; the quartiles of the distribution serve as cutoff points. Each county is colored in a shade according to the class of the respective value of the variable. Lighter shades stand for lower values and darker shades for higher values.

Figure 2: Relative Annual Number of Hospital Nights by German States

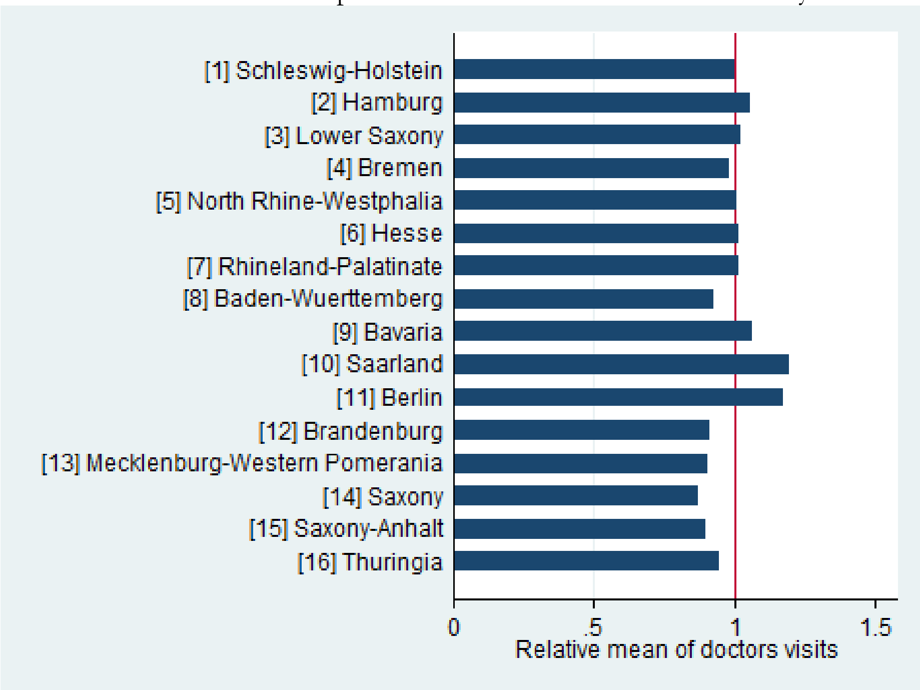
Sources: SOEP v 27, own calculations. The mean annual number of hospital nights relative to the national mean is displayed. The blue line represents the national mean.

Figure 3: Number of Outpatient Doctor Visits in the Last Three Months by German States (Frequency Weighted)



S ources: SOEP v27, own calculations. The mean number of outpatient doctor visits during the last three months is displayed. The solid black line highlights the territory of the former GDR. The values of the variable are divided into four classes; the quartiles of the distribution serve as cutoff points. Each county is colored in a shade according to the class of the respective value of the variable. Lighter shades stand for lower values and darker shades for higher values.

Figure 4: Relative Number of Outpatient Doctor Visits in Three Months by German States



Sources: SOEP v 27, own calculations. The mean number of outpatient doctor visits in the last three months relative to the national mean is displayed. The red line represents the national mean.

Table 1: Summary Statistics

Variable	Mean	Std. Dev.	Min	Max	N
Dependent Variables					
Nights spent in hospital during the last calendar year	1.51	7.61	0	280	54,362
Number of doctor visits during the last three months	2.45	3.81	0	99	54,362
A: Individual Characteristics					
Demographics					
Age	49.35	17.59	17	100	54,362
Number of children under 14	0.36	0.75	0	8	54,362
Male	0.48	0.50	0	1	54,362
Married	0.59	0.49	0	1	54,362
Education & Labor Market Participation					
Monthly gross income (*equalized household income*)	1,778	1,264	0	66,666	54,362
Not working	0.43	0.50	0	1	54,362
University degree	0.21	0.41	0	1	54,362
No vocational training	0.23	0.42	0	1	54,362
Private health insurance	0.15	0.36	0	1	54,362
Health & Health Behavior					
SF12 physical Health - generic health measure	49.23	10.10	9.93	74.94	54,362
SF12 mental Health - generic health measure	50.20	10.06	5.32	79.43	54,362
SAH good health - self-assessed (best of 5 categories)	0.09	0.28	0	1	54,362
SAH bad health - self-assessed (worst 2 of 5 categories)	0.18	0.38	0	1	54,362
No alcohol consumption	0.13	0.34	0	1	54,362
Regular alcohol consumption	0.17	0.38	0	1	54,362
Smoker	0.27	0.45	0	1	54,362
Body Mass Index (BMI)	25.96	4.60	12.59	73.05	54,362
Regular sports	0.37	0.48	0	1	54,362
B: Regional Characteristics					
Inhabitants per physician	625.27	68.85	413.30	784.20	54,362
Hospital beds per 10,000 inhabitants	61,24	8,23	15.00	85.40	54,362

Table 2: OLS Regression Models Controlling for a Rich Set of Individual-Level Covariates to Explain State Level Differences in Health Care Utilization

	Nights spent in hospital during the last year			***Number of doctor visits during the last three months***		
<u>**Covariates**</u>						
Demographics, education & labor market participation						
	Model 1a	**Model 1b**	**Model 1c**	**Model 2a**	**Model 2b**	**Model 2d**
Age	0.03592		-0.00209	0.03423 ***		-0.01358
	0.02468		*0.02493*	*0.00963*		*0.00907*
Age Squared/100	0.00414		0.00491	-0.00523		0.00574
	0.02604		*0.02710*	*0.00986*		*0.00927*
Number of Children under 14	-0.05726		-0.03815	-0.04104		0.02343
	0.05206		*0.05290*	*0.03620*		*0.03412*
Male	0.15068		0.48904 ***	-0.41662 ***		-0.20490 ***
	0.11329		*0.13401*	*0.04701*		*0.04719*
Married	-0.38797 ***		-0.31997 **	0.13005 **		0.13872 ***
	0.14389		*0.14277*	*0.05265*		*0.04894*
Monthly Income/100	-0.01116 ***		0.00540 *	-0.00435 **		0.00809 ***
	0.00354		*0.00324*	*0.00191*		*0.00181*
Not working	1.39966 ***		0.95608 ***	0.72357 ***		0.37398 ***
	0.18978		*0.16892*	*0.06451*		*0.05900*
University degree	-0.22280 **		-0.03398	-0.13608 **		0.03794
	0.10301		*0.10564*	*0.05641*		*0.05306*
No vocational training	-0.17617		-0.48356 ***	-0.01937		-0.20610 ***
	0.16716		*0.16364*	*0.06889*		*0.06315*
Private Health Insurance	-0.22161 *		-0.00400	-0.02947		0.16049 **
	0.12266		*0.12204*	*0.07420*		*0.07100*
Health & Health Behavior						
SF12 physical health		-0.13327 ***	-0.11832 ***		-0.11768 ***	-0.11789 ***
		0.00881	*0.01009*		*0.00359*	*0.00424*
SF12 mental health		-0.06484 ***	-0.06734 ***		-0.04949 ***	-0.04905 ***
		0.00917	*0.00931*		*0.00309*	*0.00308*

SAH good health		0.76768 ***	0.71300 ***		0.22309 ***	0.20719 ***
		0.08632	*0.09178*		*0.05064*	*0.05104*
SAH bad health		1.53156 ***	1.51430 ***		1.16687 ***	1.17062 ***
		0.21202	*0.20664*		*0.08755*	*0.08672*
No alcohol consumption		1.16144 ***	1.14396 ***		0.21758 ***	0.20429 ***
		0.22530	*0.23053*		*0.07425*	*0.07421*
Regular alcohol consumption		0.18841 *	-0.00535		-0.12651 ***	-0.09281 *
		0.10273	*0.10606*		*0.04846*	*0.05002*
Smoker		-0.46269 ***	-0.36480 **		-0.35065 ***	-0.29710 ***
		0.11536	*0.12204*		*0.05059*	*0.04967*
Body Mass Index (BMI)		-0.04107 ***	-0.04356 *		-0.00005	0.00475
		0.01566	*0.01765*		*0.00543*	*0.00558*
Regular sports		-0.04984	-0.01841		0.29516 ***	0.25114 ***
		0.12199	*0.11854*		*0.04143*	*0.04171*
Health care infrastructure						
Inhabitants per physician			0.01247			-0.00293
			0.08076			*0.00316*
Hospital beds per 10,000 inhabitants			-0.02185 *			-0.00048
			0.01226			*0.00429*
State fixed Effects	x	x	x	x	x	x
Year fixed Effects	x	x	x	x	x	x
Month fixed Effects	x	x	x	x	x	x
R^2	0.02350	0.06040	0.06460	0.04840	0.18480	0.18800

N=54,362; Significance levels: * = 10%, ** = 5%, *** = 1%. The robust standard errors are given in italics and clustered on the individual level. Each column yields one OLS regression model. All regression models control for month and year fixed effects. Moreover, all models include 15 state dummies. For the sake of saving space, the coefficients are not displayed since they are small in size and mostly insignificant. In Models 1a-c, between one (Model 1a) and three (1b, 1c) state dummies yield unsystematic and marginally significant effects. Only for the metropolitan region of Hamburg, there is some evidence that inpatient care utilization might be higher since the coefficient is positive and marginally significant in Models 1b and 1c. For Model 2c and outpatient care utilization, we also find largely insignificant and small state-level effects. Only Lower Saxony and Rhineland-Palatinate yield some evidence for small, positive effects. All detailed results are available upon request from the authors.