Eur J Health Econ (2012) 13:635–649
DOI 10.1007/s10198-011-0372-5

ORIGINAL PAPER

Medical guidelines, physician density, and quality of care: evidence from German SHARE data

Hendrik Jürges · Vincent Pohl

Received: 8 September 2010 / Accepted: 6 December 2011 / Published online: 28 December 2011

Abstract We use German SHARE data to study the relationship between district general practitioner density and the quality of preventive care provided to older adults. We measure physician quality of care as the degree of adherence to medical guidelines (for the management of risk factors for cardiovascular disease and the prevention of falls) as reported by patients. Contrary to theoretical expectations, we find only weak and insignificant effects of physician density on quality of care. Our results shed doubt on the notion that increasing physician supply will increase the quality of care provided in Germany's present health care system.

Keywords Physician quality of care · Physician density · Preventive care · Chronic disease management · Medical guidelines

JEL Classification I11 · I12 · J44

Introduction

Despite varied efforts to contain them, health care costs are rising in all developed countries [19]. Concerns are growing that the focus on cost saving pushes quality of care (QoC) from the health policy agenda, and that the quality of medical care stagnates or even declines in spite of further cost increases [35]. Systematic evidence on QoC is still scarce, but a growing number of studies suggest that QoC is substandard not only in the US, but in most other countries—independent of how health care is organized (e.g., [20, 33, 14]). The goal of this paper is to add evidence to this literature by studying quality of care in the German fixed fee-for-service health care system. In particular, we investigate the relationship between local primary care physician (general practitioner) density and the *process* QoC they provide. In the primary care context, process measures are preferred to outcome measures, because the latter are affected by many factors that are beyond the reach of the primary care physician—from societal factors to other levels of health care provision and their quality of care [26].

Common process measures include the use of preventive services (e.g., screening for breast cancer, influenza vaccination, advice for smokers to quit) and chronic disease management (diet and exercise counseling, lifestyle modification, regular weight checks). These services are standard subjects of medical guidelines. Over the last two decades, medical societies and other institutions have developed such evidence-based guidelines for the appropriate prevention, diagnosis and treatment of common symptoms and diseases, providing physicians with recommendations in specific situations. Guidelines are designed not only to safeguard against underuse of care or wrong care, they are also a measure of cost containment (e.g., in managed care programs), i.e., a measure to reduce overuse of medical services, for instance unnecessary tests and treatments. Thus, they take not only effectiveness but also efficiency considerations into account [11]. Since guidelines are evidence-based, representing the best available knowledge, adherence to guidelines can be interpreted as an indicator of QoC.

H. Jürges (✉)
Schumpeter School of Business and Economics,
University of Wuppertal, Rainer-Gruenter-Str. 21,
42119 Wuppertal, Germany
e-mail: juerges@uni-wuppertal.de

V. Pohl
Department of Economics, Yale University,
P.O. Box 208268, New Haven, CT 06520, USA
e-mail: vincent.pohl@yale.edu

As in many other countries, office-based physicians are the main providers of primary care in Germany. Considering the dual imperative of saving costs and providing patients with appropriate QoC, our main research question is how QoC provided by primary care physicians varies with physician supply. More physicians per capita should improve access to health care by reducing average travel time to care providers, shortening waiting lists, increasing the length of physician–patient encounters and the number of follow-up visits. Intuitively, increased time per patient should enable physicians to follow recommendations from guidelines more closely and thereby increase quality per patient. Another argument, borrowed from the monopolistic competition literature, is that increased physician density entails more competition for patients among physicians. If patients value QoC and if fees are fixed (as is the case in Germany), physicians compete by increasing quality. In equilibrium, QoC should be positively correlated with physician density. A formal exposition of this argument is given in "An illustrative model".

The result of the present study is policy relevant both in terms of efficiency and equity. Equity concerns may be raised if better QoC is provided in high physician-density areas, in particular since, in Germany, as in many other countries, high-physician density coincides with high average income. Regional inequality in the quality of medical care would then not only be an equity concern in itself but also be identified as one possible pathway explaining the SES-health gradient. In efficiency terms, finding a positive relationship between physician density and QoC would imply that health care quality can be further improved by increasing the number of health care professionals.

To our knowledge, the present paper contains the first study on the relationship between physician supply and process QoC in Germany. There are already a number of studies for Germany that include physician density (or supply) in models of physician behavior. However, they do so exclusively in relation to the hypothesis of supplier-induced demand (i.e., physician-induced overuse of health care)—using the number of doctor visits or some other measure of health care utilization per capita as dependent variables [2, 6, 28, 41, 24].[1]

Although overuse or even wrong use certainly is a quality issue in its own right, the literature has identified underuse as the main quality problem. Using data from the German sample of the 2004 Survey of Health, Ageing and Retirement in Europe (SHARE), we analyze the relationship between district general practitioner (GP) density and the degree of their adherence to medical guidelines for the prevention of falls and the management of risk factors for cardiovascular disease as reported by patients (survey respondents). We admit that this is only a narrow selection of QoC indicators. For instance, we have no indicator of the quality of acute care. However, although such a selection cannot give a comprehensive picture of the quality provided by the German health care system, it covers two main public health problems in the older population (see "Background: cardiovascular disease, falls, and medical guidelines" below).

In our data, we find no systematic relationship between physician density and QoC; the relationship is basically flat. This finding persists through a number of robustness checks and alternative specifications. Apparently, GPs in Germany do not provide better QoC when they have fewer patients or when competition increases because of higher physician density. This result contradicts the intuition described above and has important implications. It means that if the policy makers' primary aim is to contain costs while keeping health care quality at its current level, they can do so by actually *reducing* the number of physicians per capita (provided also that less physicians per capita make a health care system less expensive). However, quality problems have been identified in many if not all health care systems, so that improving quality seems to be an urgent policy goal. Our findings suggest that, if policy makers aim to improve quality, ways other than increasing the number of physicians have to be found. In line with this result, Greenberg and Greenberg [17] argue that increasing physician supply is not the right measure to treat the growing number of chronically ill older patients. Rather, to improve quality of treatment it is more appropriate to establish a disease management model in which physicians play a leading role but which also incorporates other medical professionals. A further possibility is to change incentives and introduce performance-related pay, as was recently done in the UK [4, 25]. In sum, although our result is negative, it has important implications for tackling the current problems of many health care systems.

The rest of this paper is organized as follows: in "Related literature", we give a brief overview of the previous literature and related work. "Background: cardiovascular disease, falls, and medical guidelines" provides some background on medical guidelines. "An illustrative model" shows a simple monopolistic competition model. "Data and measurements" contains a description of the data and "Regression results" shows the empirical results. "Summary and discussion" discusses the results and concludes.

Related literature

Perhaps due to the problems inherent in measuring QoC [32], the existing literature on the relationship between

[1] See McGuire [36] for a review of the vast literature on supplier-induced demand including the relationship between physician density and health care utilization.

physician density and QoC is surprisingly sparse, and the results are far from unambiguous. A couple of studies have looked at outcomes: for instance, Chen and Lowenstein [9] analyze the relationship between physician density and infant mortality—internationally and within the US. They find a strong relationship across countries but no relationship within the US, where physician density is on a relatively high level. Shi and Starfield [46] find a negative correlation between metropolitan area level physician density and total mortality rates. The correlation becomes insignificant, however, when socio-economic status is controlled for.

A common indicator of the effectiveness of primary health care is the avoidance of hospitalization for ambulatory care sensitive (ACS) conditions. These conditions do not require hospitalization when treated properly and early enough. Examples are asthma, diabetes, chronic obstructive pulmonary disease, and congestive heart failure. Laditka [29] finds that low and high physician density areas have higher hospitalization rates for ACS conditions than medium density areas (measured in density quartiles). The increase from middle to high-density areas can be interpreted as evidence in favor of the supplier-induced demand hypothesis. Laditka et al. [30] find a negative association of ACS rates with physician supply in urban but not in rural areas. In contrast, Krakauer et al. [27] find that physician supply has only a small and insignificant effect on ACS rates. Using a different outcome, Morris and Gravelle [38] find a negative relationship between physician density and body mass index (BMI), suggesting that higher physician supply improves the management of obesity.

In one of the early examples of this literature, Perrin and Valvona [40] look at the impact of increased physician supply on process quality. Distinguishing appropriate, discretionary, and inappropriate ancillary tests and treatments as measures of quality, they find small (and usually negative) effects of physician density on the frequency of ancillary tests (whether indicated or not). Based on this finding, they recommend to give physicians incentives to change their behavior instead of increasing physician supply in order to increase overall quality of health care. One of the problems of their study is that it is not entirely clear what constitutes an appropriate or inappropriate test. The study was conducted before medical guidelines emerged, and the advantage of guidelines lies in their establishing appropriateness based on evidence of clinical trials and defining exactly which tests or treatments should be conducted instead of relying on subjective treatment styles.

Although their primary aim was to study the relationship between physician advertising and the quantity and quality of physician services, some of Rizzo and Zeckhauser's [43] results also address the question asked in the present paper. Measuring the quality of physician services by the average time spent per patient, they find positive but only marginally significant effects of physician density on quality. Thus one of the important theoretical links between physician supply and QoC (longer physician–patient interactions) appears to be rather weak.

Although concerned mostly with supplier-induced demand in Norway, the analysis of Carlsen and Grytten [5] also sheds light on the relationship between physician density and QoC. Instead of some direct measure of quality, the authors use various measures of patient satisfaction, such as satisfaction with "information about diagnosis and treatments," "the physician's professional skills," or the "outcome of the treatment" as dependent variables. It seems plausible that patients are able to assess their doctor's quality to a certain extent, and that their answers reflect this assessment. Carlsen and Grytten find a positive effect of physician density on patient satisfaction, which has diminishing returns.

In terms of how QoC is measured, the works by Jencks et al. [21], McGlynn et al. [34], or Campbell et al. [4] come closest to our approach. These studies analyze care actually received in relation to the appropriate health care (standard processes) suggested by the relevant medical guidelines. Although the studies find substantive deficiencies in the quality of care, they contain no evidence on the relationship between health care system inputs such as physician density and the derived quality measures. Wennberg et al. [50] go a bit further in this respect. They generally find insufficient "effective" care (by which they mean compliance with evidence-based practice guidelines) in US regional data. Moreover, they find no relationship between the level of health care spending and effective care and thus conclude, "greater spending does not purchase the infrastructure needed to ensure compliance with the standards of practice dictated by evidence-based medicine" (pp W99–W100).

Background: cardiovascular disease, falls, and medical guidelines

In Germany, the Advisory Council for the Concerted Action in Health Care (Sachverständigenrat für die Konzertierte Aktion im Gesundheitswesen) initiated the development of medical guidelines in the mid-1990s.[2] In its 1995 expertise, the Council recommended to put the Association of the Scientific Medical Societies (Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen

[2] A detailed description of the German health care system is beyond the scope of the present paper. Interested readers are referred to, e.g., Getzen [15], Ch. 17.

Fachgesellschaften, AWMF) in charge of coordinating the design of guidelines. Its main tasks are the coordination of new guidelines, providing methodology to other involved institutions, and quality management. So far, the AWMF has collected more than 1,500 guidelines published by German medical societies. These medical guidelines are designed to support physicians' decision making in specific cases of health problems. They are developed systematically and based on scientific evidence. Notably, guidelines are not inflexible directives but provide a range of recommendations, from which the physician can deviate when appropriate [11].

The measures we use as indicators of process QoC are related to the management of risk factors for cardiovascular disease (hypertension, hypercholesterolemia, obesity, diabetes) and prevention of falls among older patients. Several guidelines have been developed in Germany for these two conditions. Our choice is driven by the fact that different guidelines in case of CVDs and falls are in fact very similar in terms of the recommended screening procedures that we analyze as process QoC.

Cardiovascular disease (CVD) was the leading cause of death in Germany in 2004. Among men aged 25–74, CVD was responsible for 33% of all deaths, whereas this number increased to 61% for men over the age of 90. The respective numbers were 28 and 66% for women [31]. Major risk factors for CVD are smoking, alcohol consumption, an unhealthy diet containing a lot of saturated fatty acids, and obesity. A recent study from Denmark, for instance, shows that obesity accounts for the largest increase in the risk of acute coronary events, even when lifestyle behaviors such as smoking are controlled for [22].

The German Cardiac Society (Deutsche Gesellschaft für Kardiologie, DGK) published a guideline [16] before the SHARE data were collected so that we can assume that the respondents' general practitioners were aware of the recommendations for the prevention of cardiovascular disease. This guideline lists a number of preventive interventions that can reduce the risk of CVD while avoiding a medicinal treatment for patients who have been diagnosed with any component of CVD or who are considered at risk of developing one of these symptoms. These interventions include smoking cessation, low-fat nutrition with high fiber content, reduction of overweight or obesity to a BMI of below 25, and physical activity. Physicians should counsel patients on changing their lifestyle behavior and monitor their progress in achieving this goal. These recommendations are reflected in the variables we use to construct our measure of recommended care: measuring patients' weight, asking about physical activity, and advising them to get regular exercise.

Falls are the primary cause of fatal and nonfatal injuries among individuals aged 65 and over [48]. About one-third of the population aged 65 and over fall at least once per year. The risk increases with age, with 80- to 89-year-olds having a 40–50% risk of falling in a given year. The risk is higher for women (relative risk 1.2–1.8). Falls are a serious health and prevention issue. About one-third of those who fall sustain an injury that results in a doctor visit or activity restrictions for at least 1 day and about 1% of all falls eventually lead to the patient's death. Even minor fall-related injuries can have a long-lasting effect on the quality of life of older patients by inducing a fear of falling, which can lead to a decline in physical and mental performance, an increased risk of falling and progressive loss of health-related quality of life [44, 49].

The German Society for General Medicine (Deutsche Gesellschaft für Allgemeinmedizin, DEGAM) published a guideline about falls and fall prevention among older patients in 2004 [18].[3] GPs should ask their elderly patients whether they have had a recent episode of falls or near falls and if so under what circumstances the fall occurred. This investigation should be the basis for a whole range of possible interventions: walk and balance training, improving visual acuity, checking and possibly changing medication, and even checking the patient's environment (especially at home) for possible causes of risk (e.g., loose rugs). Besides asking the patient about falls the physician can also assess her propensity to fall directly using several possible tests. The DEGAM guideline suggests different tests such as a "walk and count" test, which measures the patient's walking speed while she is counting backwards from 100 in steps of three, and a test involving getting up from a chair and walking. The indicators available in our data include both dimensions: the respondents indicate whether their GP asks about falls and whether their balance is checked.

Before coming to our empirical analysis, some final remarks on physician remuneration are necessary. GPs might perform each of the tasks outlined in the medical guidelines because they can bill their patients' health insurance for each task separately. This is not possible in the German fixed fee-for-service system. The fees that the Physicians' Association (Kassenärztliche Vereinigung) negotiates with health insurance companies can include several items. These fees are recorded in a scale of fees for physicians (Einheitlicher Bewertungsmaßstab, EBM). For instance, a GP can bill 8.74 Euros for an extensive

[3] Note that we are generally not able to tell whether the GP of a specific respondent was aware of the existence or development of falls prevention guidelines at the time of the interview. Data collection for SHARE wave 1 lasted from May 2004 to October 2004, while the DEGAM guidelines were available officially from July 2004 onwards. However, the publication of a guideline is only the last step of a multi-stage development process of which GPs could have been aware. The near publication of a guideline could have raised awareness for falls prevention in general.

consultation (exceeding 10 min).[4] Such a consultation would include all the tasks detailed in the guidelines for falling and cardiovascular disease. The fee scale contains no items for these checks and therefore physicians cannot bill them separately. Hence, GPs have no financial incentive to perform these services. This leaves quality competition and work ethic as possible motives for physicians to increase their adherence to guidelines when they have fewer patients.

An illustrative model

This section lays out a simple model of monopolistic competition in the market for primary medical care. In contrast to Dranove and Satterthwaite's [13] general model, physicians control only the quality of their service.[5] In the German system where physicians' associations ("Kassenärztliche Vereinigungen'') negotiate prices with health insurance companies, physicians cannot set the prices of their services individually. Hence, we assume that physicians take prices p as given and set quality q to maximize their income. Patients value quality according to the benefit function $b(q)$ and the cost of quality for physicians measured in monetary terms is $c(q)$. We assume that the benefit function depends on true quality q. This implies that patients can observe the quality provided by their medical providers. Although this is a simplification, we do not expect it to make a major difference in the present context since the quality measures we use in the empirical analysis (see discussion above) are easily observable by patients.

We make the following functional form assumptions: $b' > 0, b'' < 0, c' > 0$, and $c'' > 0$. The number of (identical) patients n that demand services from a particular physician depends on the benefit he provides and on physician density d in the area where he practices: $n(b(q),d)$. Physician density is the inverse of patients per physician so that physicians have fewer patients when d increases. The number of patients per physician increases in the benefit provided and weakly decreases in physician density: $\partial n/\partial b > 0$ and $\partial n/\partial d \leq 0$.[6] Furthermore, $\partial^2 n/\partial b^2 < 0$ since the number of patients cannot grow without bounds as benefits of medical services increase, and we assume that $\partial^2 n/(\partial b \partial d) = 0$, there are a fixed number of patients requiring medical care at any given time and patients prefer a physician who is "closer", where closeness can be interpreted in terms of location or in terms of preferences for certain treatment style. In both cases, a higher physician density implies that the "distance" between any patient and physician does not increase and that it decreases for some patients. Hence, the number of patients per physician decreases weakly in physician density. Although patients differ in their location, they do not differ from the viewpoint of the physician.

Since all patients are identical, physician's income equals the number of patients times profit per patient:

$$y = n(b(q), d)[p - c(q)] \tag{1}$$

The participation constraint for physicians requires that he at least breaks even for every patient he treats, hence

$$p - c(q) \geq 0 \tag{2}$$

In addition, patients have an outside option such as going to a hospital or not receiving treatment at all, which provides them with benefit $\bar{b}$ so that the participation constraint for patients is

$$b(q) \geq \bar{b} \tag{3}$$

Maximizing income y with respect to quality q subject to the two constraints (Eq. 2) and (Eq. 3) yields the following first-order conditions:

$$\frac{\partial n}{\partial b} b'[p - c(q)] - nc' - \lambda_1 c' + \lambda_2 b' = 0 \tag{4}$$

$$\lambda_1 [p - c(q)] = 0 \tag{5}$$

$$\lambda_2 [b(q) - \bar{b}] = 0 \tag{6}$$

We are interested in the change of the optimal quality level as physician density varies. Assuming an interior solution with $\lambda_1 = \lambda_2 = 0$, this change equals

$$\frac{\partial q}{\partial d} = \frac{\frac{\partial n}{\partial d} c'}{\left(\frac{\partial^2 n}{\partial b^2} b'^2 + \frac{\partial n}{\partial b} b''\right)[p - c(q)] - 2\frac{\partial n}{\partial b} b'c' - nc''} \geq 0 \tag{7}$$

since the numerator is weakly negative and the denominator is strictly negative by the functional form assumptions made above. Hence, quality weakly increases with physician density when there is an interior solution.

If only the constraint in Eq. 3 binds the change of optimal quality in response to a change in physician density becomes

$$\frac{\partial q}{\partial d} = \frac{\frac{\partial n}{\partial d} c'}{\left(\frac{\partial^2 n}{\partial b^2} b'^2 + \frac{\partial n}{\partial b} b''\right)[p - c(q)] - 2\frac{\partial n}{\partial b} b'c' - nc'' + \lambda_2 b''} \geq 0 \tag{8}$$

[4] This amount applies to patients with public health insurance. When a GP treats privately insured patients, he can bill 20.10 Euros for the same service.

[5] In the model in Dranove and Satterthwaite's [13] review, physician choose an optimal price as well as one or more non-price attributes of their services.

[6] In a dynamic model, we could assume that the number of patients per provider also depends on past patient numbers. In other words, patients incur switching costs when changing providers and are therefore more likely to stay with a provider. This would decrease only the effect of density on number of patients per provider (in absolute value), but not make it positive (higher density can never lead to more patients per provider). In the following, we consider only a static model of quality choice and abstract from switching costs.

for positive changes in physician density. Since λ_2 is strictly positive and b'' is strictly negative by assumption, this change is strictly larger than in Eq. 7. Thus, quality changes more in response to variations in physician density. When physicians provide their patients with quality such that patient benefit just equals the value of their outside option there is more scope for improvement in quality when the average number of patients per physician decreases.

On the other hand, if only the constraint in Eq. 2 is binding, the first-order condition reduces to $-nc' - \lambda_1 c' = 0$, which is impossible since n is positive. In this case, the constraint given in Eq. 3 must also bind so that the first-order condition becomes

$$-nc' - \lambda_1 c' + \lambda_2 b = 0 \tag{9}$$

This implies that whenever the physician's break-even constraint is binding, he provides his patients with a quality level that just satisfies their participation constraint. Using the first-order condition (Eq. 9), the change in quality in response to variation in physician density becomes

$$\frac{\partial q}{\partial d} = \frac{\frac{\partial n}{\partial d} c'}{-\frac{\partial n}{\partial b} b'c' - nc'' - \lambda_1 c'' + \lambda_2 b''} \geq 0 \tag{10}$$

since all terms in the denominator are strictly negative and the numerator is less than zero. In sum, we conclude that quality provided by physicians weakly increases with physician density.

Data and measurements

The data used in this study combine survey data on physicians' adherence to medical guidelines and relevant individual characteristics with indicators on the regional level, specifically on the district (Kreis) level. The former are drawn from the German subsample of the first wave of the Survey of Health, Aging, and Retirement in Europe (SHARE), collected in 2004. SHARE was designed to provide comparable multi-disciplinary data across European countries. It is modeled closely on the US Health and Retirement Study (HRS) and contains rich information on the respondents' health, saving and retirement decisions, demographic characteristics and many other dimensions. Representative samples of the non-institutionalised population aged 50+ in each country were interviewed using computerised face-to-face questionnaires and self-completion paper questionnaires. Samples were drawn from population registries, or from multi-stage probability sampling, in by now 14 European countries and Israel. Specific details are provided elsewhere.[7] District level data stem from the 2004 regional database (INKAR) of the Federal Office for Building and Regional Planning (BBR), which is a rich source of statistical information on the district level in Germany (for the year 2002). Germany currently has 439 districts. SHARE respondents have been sampled from a subset of 100 randomly selected districts.

Quality of care

Questions about GPs' adherence to guidelines (i.e., process QoC) were asked in the self-completion part of SHARE. Three items are related to the management of CVD-related chronic conditions such as hypertension, hypercholesteremia, diabetes, and obesity: respondents indicate how often their GP asks them about physical activity, tells them to get regular exercise, and checks their weight ("never," "at some visits," or "at every visit"). Two items are related to falls and the risk of falling: how often the GP asks respondents about falling down and how often the GP checks their balance or the way they walk. These items mirror the recommendations from the guidelines described above.

For our empirical analysis, we summarized the information given by the respondents as follows. As a first step, we have collapsed "always" and "at some visits" into a single category. We do this because it is unclear if, say, asking about physical activity at every visit rather than at some visits is that important in cases of patients who visit their GP many times a year. Also, the distinction between "always" and "at some visits" might be blurred due to recall bias. Thus we recoded the respondents' answers so that a value of 1 indicates that the respondents' GP has performed these checks at least at some visits (there is no reference period).

In the second step we computed the *percentage of recommended care* received by each respondent eligible for that type of care. Note that only eligible respondents are included in our analytical samples. Following guidelines, all respondents aged 65 are classified as being eligible for falls prevention ($N = 772$). If the GP has neither asked about falls nor checked the patient's balance, the respondent has received 0% of recommended care. If the GP has either asked about falls or checked balance (but not both), the respondent has received 50% of recommended care, and if the GP has asked about falls and checked balance, the respondent has received 100% of recommended care. Overall, we find that patients eligible for falls related care receive on average 42% of recommended care. In 49% of the cases, respondents said they received neither of the two measures, 18% said they received only one of the two measures and 33% said they received both measures.

Each respondent who reports to have been diagnosed with, or reports to take medication for, hypertension,

[7] For a detailed description of SHARE and data access see http://www.share-project.org.

hypercholesteremia, diabetes, or heart disease, or who is overweight (BMI > 25 kg/m^2) is classified as being eligible for the three CVD-related types of care measured in SHARE, which are lifestyle related ($N = 1{,}135$). With three measures, respondents can have received 0, 33, 67, or 100% of recommended care. We find that, on average, patients eligible for CVD-related care received 58% of recommended care. Roughly 20% of the respondents eligible for care said they received none, one, or two of these measures, respectively. The remaining 37% received all three.

Substantial regional differences in QoC have been identified across US states [21, 7], but there is little evidence for Germany. Figure 1 shows the cumulative distribution of district means for our two QoC indicators. The district median of the percentage of received recommended care is 58% for CVD risks and 42% for falls. The range between the first and third quartiles is from 50 to 71 for CVD risks and 25 to 54 for falls. Hence, there is substantial variation in process QoC between districts. In particular, QoC for fall prevention is unsatisfactory in a majority of districts.

Since we use data only for Germany in our later analyses, an international comparison of QoC is useful to place our results into a broader context. Figure 2 shows average QoC measured as the average percentage of recommended care received in those eight SHARE countries where we also have primary physician density (from the European Health for All Database, available at http://www.euro.who.int/hfadb). According to the respondents' answers, the best QoC with respect to the management of CVD-related chronic conditions is delivered in France (respondents received 68% of recommended care). The worst QoC is delivered in the Netherlands, where respondents received about 40% of recommended care. In Germany, we find an average performance with respondents receiving 60% of recommended care. The overall performance with respect to the prevention of falls is generally worse than with respect to the management of CVD-related chronic conditions. Respondents received on average between 22% (again in the Netherlands) and 46% (France) of recommended care. In international comparison, Germany performs better than average (43%).[8] What is striking in Fig. 2 is the positive association between country-level primary physician density and QoC (statistically significant only for falls). Of course this association has to be interpreted with caution because different countries have organized primary care differently, with doctors being involved in primary care to varying degrees. Omitted variables bias is a much more likely problem on the country level. We thus prefer to continue our analysis within a country and health-care system.

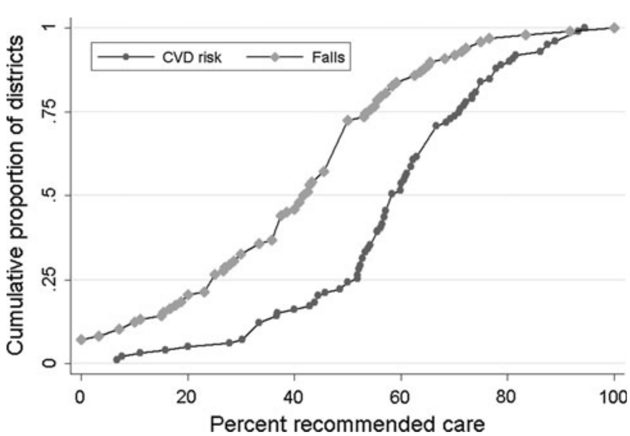

Fig. 1 Cumulative regional distribution of quality of care (QoC) indicators

Respondent-level covariates

In our regression analysis for Germany, we use the following individual-level control variables: age, sex, marital status, education, and self-perceived general health. Moreover, we use specific individual health variables related to the two health problems addressed in this paper: first, we computed a summary index of CVD-risk factors by simply counting the number of risk factors we use to define eligibility for care: hypertension, hypercholesteremia, diabetes, overweight (BMI > 25 kg/m^2), and already having been diagnosed with heart disease. A separate indicator is included for obesity (BMI > 30 kg/m^2). The risk of falling is assessed by the sum of three indicators: whether the respondent has experienced a fall in the past, suffers from fear of falling, or suffers from dizziness, faints or blackouts. We expect to see an impact of these risk indicators on the proportion of recommended care received by respondents. GPs should provide better health care to those who need it more.

We also include number of visits to a GP in the last 12 months in our regressions (in logarithms—which effectively eliminates all respondents with zero GP visits in the last year from the estimation). Seeing a physician more often increases the likelihood of having the tasks under consideration performed at least at some visits. This is an artifact of the way the questions were asked in the questionnaire and our coding of the answers. In the question, no time horizon was specified. In particular, respondents were not asked, for instance, how many times their GP had checked their weight, etc. during the past year. Finally, we include a dummy variable for having private health

[8] Note that primary care physician density in Germany (102.4) is substantially larger in the European Health for All Database than in the INKAR database (58.3—if one-third of all internists are included). We do not know where this difference comes from, as the data sources of the European Health for All Database for specific indicators are not documented.

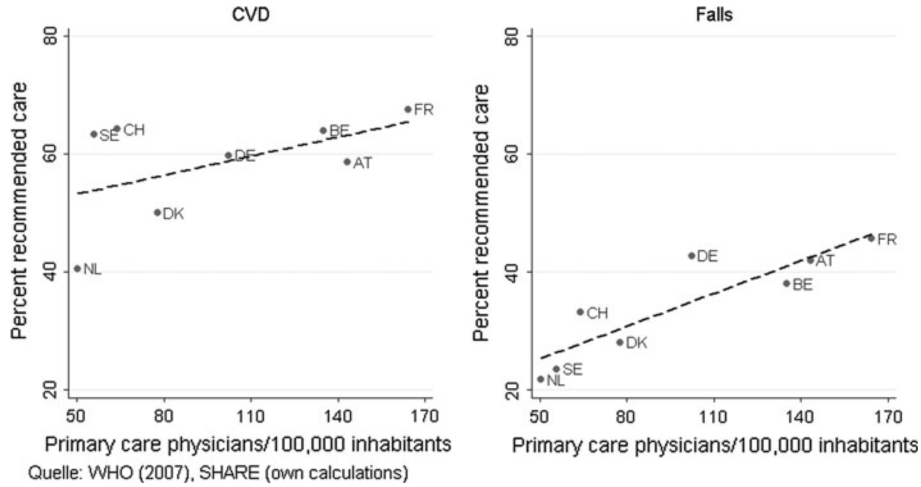

Fig. 2 Cross-country comparison of QoC (*GR* Greece, *FR* France, *CH* Switzerland, *BE* Belgium, *SE* Sweden, *ES* Spain, *AT* Austria, *DE* Germany, *DK* Denmark, *IT* Italy, *NL* Netherlands)

insurance. Having private health insurance is likely to increase the probability of GPs performing certain checks since they receive higher fees for the privately insured and might want to provide them with higher quality care in order to retain them as patients.[9]

Table 1 contains summary statistics for these variables, separately for our two analytical samples. Members of the sample used for the analysis of falls are, on average, older than members of the CVD sample. This is not surprising since the risk of falls increases with age whereas symptoms such as obesity and hypertension also occur in younger patients. Since age and health are negatively correlated, members of the falls sample are also more likely to have poor self-rated health and visit their general practitioner more often. In the CVD sample, the average number of CVD risk factors is about 2 out of 5; a fifth of this sample are obese. The average number of fall-related symptoms is about 0.4 out of 3 in the falls sample with a minimum of 0. That is, there are individuals in the sample who have not experienced fall-related symptoms, but are included in the sample due to their age. The remaining individual covariates vary little between the two samples.

District-level covariates

We derived information on district physician density, measured as the number of physicians per 100,000 inhabitants from the INKAR database.[10] Since SHARE respondents are asked explicitly about services performed by their GPs, we use the number of GPs per 100,000 inhabitants. In Germany, office-based internists often fulfill the role of GPs, we thus added to this number 30% of the number of internists per 100,000 inhabitants.[11] In the districts covered by SHARE, GP density ranges from 43.3 to 86.7 with a median of 58.4 and an interquartile range of 9.7.[12] Although more than 50% of all districts have between 54 and 64 GPs per 100,000 inhabitants, there is sufficient regional variation outside this range for an informative empirical analysis.

To test the robustness of our results, we also use two alternative measures of physician density. One is the number of GPs per 100,000 inhabitants aged 50 and over (median = 161, IQR = 33). Older patients typically see their doctor more often and require more care. The GPs' effective time per patient might thus be better approximated by the ratio of doctors to older patients. We also use a space-based density measure, namely the number of GPs per 100 square kilometer. The distribution is highly skewed (minimum = 2.7, median = 16.3, maximum = 323), and we use the natural logarithm in our regressions.

In addition to physician density, we use a measure for the regional age composition of GPs. In particular, we include the fraction of GPs over the age of 60 (median 14.2, interquartile range 5.3).[13] Older physicians might be less familiar with guidelines (or be less willing to follow guidelines) than younger physicians, so that we expect a negative correlation between district QoC and the proportion of GPs who are older than 60.

Finally, we also include a dummy variable for Eastern German districts, where 22% of our sample lives. There are

[9] See [24] for the relation between insurance status and supplier induced demand.

[10] The number of physicians includes those providing ambulatory services under contract of the regional doctors' association (Kassenärztliche Vereinigung). Such a contract is a necessary condition to treat patients insured in the German statutory health care system. Physicians who treat exclusively privately insured patients (less than 5% of all physicians providing ambulatory services) are thus not included here.

[11] This is somewhat arbitrary, but our results are not sensitive to including or excluding internists into our measure of GP density.

[12] The range across all German districts is from 38 to 90, so the SHARE districts are reasonably representative in this respect.

[13] These data were drawn from the physician database of the Kassenärztliche Bundesvereinigung (KBV).

Table 1 Sample description (averages and proportions)

Variable	CVD sample				Falls sample			
	Mean	Standard deviation	Min	Max	Mean	Standard deviation	Min	Max
Age	65.4	9.3	50	94	71.4	7.4	50	97
Male	0.48		0	1	0.45		0	1
Married	0.75		0	1	0.70		0	1
Low education	0.18		0	1	0.23		0	1
High education	0.24		0	1	0.22		0	1
Poor self-rated health	0.52		0	1	0.59		0	1
Number of GP visits	6.6	9.1	1	98	7.6	10.6	1	98
Private health insurance	0.07		0	1	0.06		0	1
Number of CVD risk factors	1.98	0.99	1	5				
Obese	0.22		0	1				
Number of fall-related symptoms					0.38	0.65	0	3
GPs per 100,000 inhabitants	59.5	8.5	43.3	86.7	60.4	8.6	43.3	86.7
GPs per 100,000 inhabitants 50+	164.2	26.6	114.9	251.9	166.1	27.0	114.9	251. 9
GPs per 100 square km	46.1	61.5	2.7	322.6	47.3	60.4	2.7	322.6
Proportion GPs 60+	16.3	6.6	0	38.4	16.1	6.6	0	38.4
Eastern Germany	0.21		0	1	0.23		0	1
N	1,135				772			

several reasons why GPs who were educated in the former GDR might provide better quality (as measured in our study) than GPs who were educated in the West. GPs (working either in their own practices or in polyclinics) formed the backbone of the GDR health care system. They had better opportunities for continuing education than their colleagues in the West. Prevention played a bigger role than in the West, and medical "guidelines" have been known for a long time. The first guidelines were developed as early as 1977, while practical medicine has been less dependent on cutting-edge medical technology (due to a lack of resources). Not accounting for an "East Germany" effect could downward bias our results on the effect of physician density, because physician density in the East is particularly low, especially in rural areas.

Regression results

We now study the effect of GP density on quality of care by regression analysis, controlling for individual and district characteristics as described in the preceding section. The regression analysis will also inform us about the importance of individual correlates of quality of care received by patients in Germany. In order to account for between district heterogeneity in unobserved characteristics, we use a random effects model to estimate the impact of physician density on adherence to medical guidelines as a measure for QoC. We include district means for all individual level variables so that the estimated random effects are orthogonal to the observed explanatory variables by construction. Formally, the econometric specification can be written as

$$y_{ik} = \alpha Z_k + \beta \bar{X}_k + \gamma X_{ik} + u_k + \varepsilon_{ik} \tag{11}$$

where y_{ik} is the percentage of recommended care received by individual i in district k, Z_k is a vector of characteristics of district k, in particular GP density. $\bar{X}_k$ is a vector of within-district averages of the individual characteristics X_{ik}. u_k reflects district level unobserved heterogeneity (the "random" effect) and ε_{ik} is an i.i.d. error term for individual i in district k.

Basic specification

Table 2 shows the basic regression results. We use a standard linear random effects estimators, so that regression parameters measure the effect of a one unit increase in X or Z on the percentage of recommended care received by individual i. Parameters β (belonging to within-district averages) are not shown for the sake of brevity. We report three different specifications per QoC indicator, each with a different GP density measure.[14]

[14] Note that coefficients for individual-level variables are identical across different specifications. These variables are orthogonal to district level variables because district-level averages of all individual variables are included in the equations.

For the number of GPs per 100,000 inhabitants, we find very small *negative* coefficients on the percentage of recommended care for CVD risks and falls received by patients. Both coefficients are far from significant. Changing our density measure to the number of GPs per 100,000 inhabitants aged 50 and over does not change this basic result. The coefficient for falls prevention turns positive, but the effect size is practically zero. For instance, the bottom quintile of districts has an average density of 129 and the top quintile has an average density of 203. Going from bottom to top would increase received recommended care by roughly 3 percentage points. This is less than the difference between men and women. Finally, for our space related measure "number of GPs per 100 square km," of which we took the logarithm, we also find insignificant effects of different signs.

Overall, our results suggest that physician density and QoC as measured in our study are virtually unrelated. When GPs have to treat a larger (or smaller) number of patients they do not appear to decrease (or increase) the quality of the care they provide. We give a detailed discussion of some possible explanations for our finding, which we are not able to test in our data, in the concluding section. Before, we address those explanations that we can test.

Robustness checks

One reason for our finding of no relationship between physician density and QoC could be that quality of care is fairly high in high physician density areas and cannot easily be improved upon if the number of patients per doctor increases further. Such improvements would then be found only in low-density areas. In other words, the relationship between physician density and QoC might be nonlinear. In order to test this conjecture, we estimated the density–quality relationship as a piecewise linear function with one kink at the median of the explanatory variable. The results are shown in Table 3. The coefficient for "GP density below median" reflects the slope estimated for all districts with below median physician density (computed separately for each measure and QoC indicator), the coefficient for "GP density above median" reflects the *difference* in slopes between the high- and low-density groups. For GPs per 100,000 inhabitants, for instance, we find a negative slope among low-density *and* high-density districts, but the slope is smaller in absolute terms among the high-density districts. Overall, the alternative functional form specification contains no evidence in favor of the assumption that a positive GP density effect is relevant only in low-density areas. If anything we find a negative relationship. Moreover, increasing the number of splines does not change that result.

One potential concern that needs to be addressed is endogeneity. Using a random effects model accounts for unobserved district level heterogeneity but cannot deal with possible endogeneity of physician density. Endogeneity might be present if physicians systematically prefer opening a practice in districts where quality is already above or below average. The former could be the case if physicians think they have to work less when the population is already well served. On the other hand, the latter case could occur when physicians are led by ethical considerations to improve the quality of health care in regions where it is currently low. Both potential sources of endogeneity are limited by government regulation in Germany since physicians are not entirely free to choose where to locate if they want to obtain accreditation at the public health insurance system ("Kassenzulassung"). Potential endogeneity of physician supply has received much attention in the induced demand literature.[15] Since these studies deal explicitly with the interrelation of supply and demand, simultaneity is a big concern. In the present paper, however, it is less clear if physician density is endogenous.

Although we believe that endogeneity should not matter much in the present case, we perform instrumental variable (IV) regressions to address this concern. As instruments, we use district per capita income (in logs), presence of a medical school, and share of population aged 65 and older, all measured at the district level. For these variables to be valid instruments they have to be correlated with physician density and they should not affect QoC (measured as adherence to medical guidelines) other than through physician density (this is our exclusion restriction). Correlation with the potentially endogenous regressor physician density is plausible, because our instruments can be interpreted as supply related variables. Physicians are assumed to prefer high-income areas mainly because it is generally nicer to live there. Furthermore, the proportion of privately insured patients (who pay higher fees) is also higher in high-income areas. An older population requires more care, and physicians therefore expect to have more patients in these districts. Finally, it seems plausible that physicians want to stay in an area where they have already spent a substantial portion of their lives while attending medical school.

Table 3 summarizes the results of our IV approach. Instrumental variables are always subject to close scrutiny. First, the instruments must not be weak, i.e., they must be correlated sufficiently with the endogenous regressor. An established rule-of-thumb criterion for good instruments is an F-statistic larger than 10 in a test of joint significance in the first stage regressions [47]. On that account, our instruments perform well. Another concern about

[15] See Dranove and Wehner [12] for a critique of the use of two-stage least squares in SID studies.

Table 2 Random effects regressions of quality of care (QoC) indicators on district and individual characteristics, basic specification

	CVD risk			Falls		
GP density	−0.069 (0.235)			−0.134 (0.235)		
GPs per 100,000 inhabitants 50+		−0.001 (0.076)			0.042 (0.078)	
Log # GPs per 100 square km			−0.153 (1.910)			1.325 (1.978)
Proportion of GPs aged 60+	0.719 (0.373)***	0.722 (0.376)***	0.712 (0.373)***	0.377 (0.363)	0.391 (0.364)	0.376 (0.355)
Eastern Germany	7.071 (6.205)	7.119 (6.240)	8.102 (6.451)	1.507 (5.992)	1.454 (5.997)	3.829 (6.105)
Age	0.321 (0.128)**			0.872 (0.229)*		
Male	4.966 (2.239)**			5.545 (3.421)		
Married	1.815 (2.577)			−7.056 (3.733)***		
Low education	−2.622 (3.103)			4.503 (4.210)		
High education	−5.489 (2.721)**			0.421 (4.151)		
Poor self-rated health	1.976 (2.428)			−1.493 (3.642)		
Log number of GP visits	4.002 (1.367)*			6.535 (1.948)*		
Private health insurance	7.341 (4.442)***			8.573 (7.290)		
Number of CVD-related conditions	3.354 (1.203)*					
Obesity	12.251 (2.698)*					
Number of falls-related symptoms				12.251 (2.662)*		
Observations	1,135	1,135	1,135	772	772	772
Number of districts	99	99	99	98	98	98
Rho	0.15	0.15	0.15	0.07	0.07	0.07
Between-R2	0.13	0.13	0.14	0.16	0.16	0.17
Within-R2	0.08	0.08	0.08	0.11	0.11	0.11

Standard errors in parentheses. Coefficients for individual level variables are the same for all three specifications. District means of individual variables were included in the regressions but the estimated parameters are not shown

*** Significant at 10%; ** significant at 5%; * significant at 1%

instrumental variables is that the identifying (exclusion) restrictions may not hold, i.e., that the instruments do not satisfy the orthogonality conditions. In the context of our overidentified model (we have three instruments for one endogenous regressor) we are able to test the overidentifying restrictions—using a Sargan-Hansen-test—to provide some evidence of the instruments' validity. The overidentifying test does not reject the null hypothesis in any of our models. Thus, we have some confidence in our identifying assumptions.

Coming to the actual results, we note that using IV does not lead to results that are qualitatively different from OLS regressions. Our IV estimates are fairly imprecise due to large standard errors, so that none of the estimated effects is statistically different from zero. It is thus not surprising that a Hausman-test does not reject the null hypothesis of no endogeneity (detailed results not shown). In sum, endogeneity of the regressor does not seem to be a big concern in our analysis.

Another concern about our basic specification is that our dependent variable is a fractional response variable. A linear function is not the best way to model such variables, for instance, because predicted values are not bounded between 0 and 1. However, this problem can be solved fairly easily by estimating a generalized linear model (GLM), e.g., using a Bernoulli distribution with a logit link function [39]. In Table 3 we report marginal effects (evaluated at the median of the explanatory variables) from such a model. Note that this is no random effects model. Dependencies within districts are dealt with only by estimating cluster-corrected standard errors. We find that most estimated marginal effects are actually very similar to their linear regression counterparts. Again, none of the estimated effects is statistically different from zero.

Table 3 Random effects regressions of QoC indicators on GP density, robustness checks

	CVD risk			Falls		
	GP per 100,000	GP per 100,000 aged 50 and over	log GPs per 100 square km	GP per 100,000	GP per 100,000 aged 50 and over	log GPs per 100 square km
Spline function						
GP density below median	−0.399 (0.462)	−0.088 (0.161)	−10.257 (5.747)***	0.286 (0.490)	0.142 (0.179)	−9.481 (6.062)
GP density above median	0.154 (0.356)	0.058 (0.121)	3.487 (2.717)	−0.408 (0.367)	−0.014 (0.120)	5.071 (2.749)***
Instrumental variable estimation						
GP density	0.579 (1.098)	0.036 (0.188)	−0.499 (2.849)	−0.016 (0.702)	0.242 (0.198)	−0.028 (3.097)
Sargan-Hansen statistic	2.850	3.279	3.275	3.511	1.873	3.765
P-value	0.241	0.194	0.194	0.173	0.392	0.152
First stage F-statistic	12.680	53.913	309.859	45.139	68.421	241.487
P-value	0.000	0.000	0.000	0.000	0.000	0.000
GLM (marginal effects)						
GP density	−0.059 (0.215)	0.012 (0.080)	0.261 (1.622)	−0.109 (0.174)	0.048 (0.067)	2.142 (1.880)

Standard errors in parentheses; all individual and district level characteristic as listed in Table 3 are controlled for. Instruments: log per-capita income, share of population 65+, presence of medical school. Marginal effects evaluated at the median of the explanatory variables

*** Significant at 10%, ** significant at 5%, * significant at 1%

Ancillary results

We now discuss some of the ancillary results listed in Table 2. Contrary to our expectations, we find a slightly positive (and in the case of CVD-risk management significant) effect of the proportion of GPs older than 60 on quality of care. One possible explanation could be the fact that our sample consists of patients aged 50 and over who may be more appropriately treated by physicians of similar age. The relation between patient's and physician's age is the subject of some studies in the medical literature [37, 45]. As expected, adherence to guidelines is stronger in East German states, especially with respect to CVD risks, but the difference to states in the West is not significant.

Since older patients are at higher risk of falling, a positive relationship between age and recommended care should be expected.[16] We actually find a fairly large effect: receiving recommended care rises by about 8.7% in each decade of life, even controlling for falls-related symptoms. The incidence of cardiovascular disease also increases with age, and we also find a positive effect of age on receiving recommended care independent of risk factors. Non-linear specifications have not improved the fit of the regressions.

Men receive about five percentage points more recommended care than women, and the difference is statistically significant for CVD-related care. This result is somewhat puzzling for falls since women are more likely to experience falls at all ages. Since we use survey data it is possible that the result is due to systematic bias in answering the questions on QoC. Men could tend to remember better or to overestimate the tasks their GP performs, for instance. Although this cannot be excluded a priori, it does not seem to be very likely either. This leaves us with the conclusion that women receive less recommended care than men. The relation of patients' and physicians' sex could be one possible interpretation for this fact. Studies have found that women receive better care when being treated by female doctors [10].[17] However, since we do not have data on the sex of GPs we cannot explore this explanation any further.

Based on evidence for other types of preventive medical care [23] on the impact of marital status and education on health care utilization, we expected that married and better-educated patients receive higher quality of care but this is not confirmed by our data. Rather, married patients receive less recommended care related to the prevention of falls. It is even more puzzling to observe worse CVD-related care among the high than among the medium educated respondents (which are the reference group).

Poor *general* health (conditional on specific risk factors) has an unsystematic and insignificant impact on QoC. In contrast, the *specific* health variables have highly significant and large positive effects. For instance, each additional CVD-related condition (hypertension, hypercholesterolemia,

[16] Whereas about one-third of people aged 65 fall at least once a year, this rate is 40–50% for the age group 80–89 [18].

[17] Interestingly, there is no evidence for racial differences in QoC related to the physician's race [8].

overweight, diabetes, and already suffering from heart disease) increases the received recommended care by about 3.3 percentage points. Obesity adds another 12.3 percentage points. Similarly, physicians are significantly more likely to ask about falls and check a patient's balance if the patient has already experienced a fall, is concerned about falling or suffers from faints or blackouts. Each of these symptoms adds 12.3 percentage points. Hence, patients who are more dependent on preventive care are also substantially more likely to receive recommended care. These findings also give us some confidence that our QoC measures do contain useful information about what actually happens in the respondents' doctor's office.

The coefficient on the logarithm of number of GP visits in the last year is positive and highly significant. As discussed earlier, this is most likely an artifact of our study design. Having a private health insurance increases received recommended care by eight percentage points (the difference is significant only for CVD-related care). We expected better care for the privately insured, because fees for services to the privately insured are more than twice as high as fees to the statutorily insured. That fact that the effect is only weakly or insignificant can perhaps be explained by the fact that less than 10% of the respondents in the sample have private health insurance. Thus estimates are fairly imprecise.

Summary and discussion

The relationship between physician supply and QoC has recently sparked a debate about the required physician workforce in the US. The Association of American Medical Colleges called for an increase of medical school admissions by 30% annually to meet the increasing demand for health care of an aging population [1]. In this paper, we study whether the implicit claim on which such calls are based, namely that more doctors improve the quality of health care, is corroborated by empirical evidence.

To this end, we studied the relationship between district GP density and the process quality of medical care they provide using German SHARE data. We measure physician quality as the degree of adherence to medical guidelines (for the prevention and management of cardiovascular disease and falls) as reported by patients (respondents). Contrary to theoretical expectations, we find virtually no effect of physician density on QoC. As our data also show, QoC is substandard in Germany. Patients receive about 60% of recommended care related to CVD prevention and management and 40% of recommended care related to the prevention of falls. Comparable data from other European (SHARE) countries show that it is possible to provide better care to the older population. However, a major limitation of our study is that these results are based on a very narrow selection of indicators to characterize QoC. One should thus be careful when generalizing these results and more research is needed to corroborate our main findings.

Given these limitations, we provide a series of successful robustness checks (different density measures, different functional forms, different estimation methods), there are still some potential explanations for our main result that we cannot test easily. First, the numbers of observations per district in our working samples are fairly small (about 8–12), so that average district quality might be measured imprecisely. Recall error in our dependent variables could contribute to this measurement error. Using administrative instead of survey data could alleviate this problem.[18] However, this would not be feasible in the present context because in the German fixed fee-for-service system, the services that we use as QoC indicators are not recorded in claims data. Furthermore, obtaining medical records from physicians or health insurers for general sample survey members would be prohibitively costly. Second, physician density on the district level might be too coarse to capture the behaviorally relevant local physician density (which might be on the town or even neighborhood level). Such measurement error in explanatory variable potentially biases the coefficient of interest towards zero.

Another reason for our finding that physician density and QoC are unrelated could be the organizational structure of doctors' practices in some districts. When physician density is low, and each GP has more patients, this could actually lead to better adherence to medical guidelines if GPs believe that following the guidelines can reduce the time spent on each patient. Using guidelines may pay off for GPs when they have (too) many patients. Without additional data on the physician level, however, we can only speculate about this mechanism counteracting our basic model.

Furthermore, it is not clear whether adherence to guidelines is an important way to attract patients, as suggested by our model. Physician density may not affect quality of care simply because there is no response to quality by consumers. However, the positive (and in case of CVD weakly significant) coefficient for private health insurance suggests that adherence to guidelines could be a strategy to attract patients. Of course, this does not suffice to reject the notion that physician supply does not affect the

[18] Raina et al. [42] compare patients' self-reported utilization of medical services with administrative data and find that patients recall correctly if they had contact with health care providers, but that there is substantial under- and over-reporting in volume utilization measures.

quality of care indicators used in our study because they are not important to patients or physicians. Alternative measures of physician performance, such as investment in high-tech office equipment (measuring professional goal) or physician hours (measuring patient orientation), could help distinguishing alternative explanations—although it is unclear how these could be reliably assessed by interviewing patients.

Despite potential data problems, our study clearly shows that there is room for improvement in QoC and the question remains how this can be achieved. If the main result of our study—the lack of any relationship between physician density and adherence to guideline—is to be believed, this has important implications in two respects. First, non-adherence to guidelines is most probably not a matter of lack of time or competitive pressure. Otherwise, doctors in high-density areas (who have more time per patient or higher competitive pressure) would perform better. Rather, one might believe in a lack of willingness to follow guidelines if physicians are unhappy with what is sometimes called "recipe medicine," and which threatens their autonomy. Changing the reimbursement system to include quality-related or guideline-based fees (i.e., performance-related pay) is one possibility to improve adherence to guidelines.

Second, our finding that GP supply and QoC are not related implies that increasing the supply of physicians above the current level will not improve the quality provided by German GPs and neither is reducing supply likely to harm patients. Such results are in line with those reported for the US: effective care is insufficient (i.e., compliance with evidence-based practice guidelines), and there is hardly any relationship between the level of health care spending and effective care [50]. In order to improve compliance with the standards of practice dictated by evidence-based medicine, Wennberg et al. suggest changing the organizational structure of medical care (in the US) by strengthening staff-model or group-model HMOs rather than just spending more money on health care. Implementing this kind of organizational change in Germany would mean going a long way and completely changing the way health care is organized today. In this spirit, our study simply gives another reason for substantive health care reform besides cost cutting: improved quality.

Acknowledgments We would like to thank Thomas Kopetsch, KBV, for providing district-level data on the age structure of physicians, and Andy Epstein, Frank Sloan, and participants at the 2008 Society of Labor Economists annual meeting for helpful discussions. This paper uses data from release 2.0.1 of SHARE 2004. The SHARE data collection has been funded primarily by the European Commission through the 5th framework program (project QLK6-CT-2001-00360 in the thematic program Quality of Life). Additional funding came from the US National Institute on Ageing (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, Y1-AG-4553-01 and OGHA 04-064). Further support by the European Commission through the 6th framework program (projects SHARE-I3, RII-CT-2006-062193, and COMPARE, CIT5-CT-2005-028857) is gratefully acknowledged. For methodological details see Börsch-Supan and Jürges [3].

References

1. Association of American Medical Colleges: AAMC Statement on the Physician Workforce. http://www.aamc.org/workforce/workforceposition.pdf (2006)
2. Breyer, F.: Die Nachfrage nach medizinischen Leistungen: Eine empirische Analyse von Daten aus der Gesetzlichen Krankenversicherung. Springer, Berlin (1984)
3. Börsch-Supan, A., Jürges, H.: The Survey of Health. Aging and Retirement in Europe. Methodology. MEA, Mannheim (2005)
4. Campbell, S., Reeves, D., Kontopantelis, E., Middleton, E., Sibbald, B., Roland, M.: Quality of primary care in England with the introduction of pay for performance. N. Engl. J. Med. **357**, 181–190 (2007)
5. Carlsen, F., Grytten, J.: Consumer satisfaction and supplier induced demand. J. Health Econ. **19**, 731–753 (2000)
6. Cassel, D., Wilke, T.: Das Saysche Gesetz im Gesundheitswesen: Schafft sich das ärztliche Leistungsangebot seine eigene Nachfrage? Zeitschrift für Gesundheitswissenschaften **9**, 331–348 (2001)
7. Chandra A, Skinner J: Geography and racial health disparities. NBER working paper 9513 (2003)
8. Chen, J., Rathore, S., Radford, M.J., Wang, Y., Krumholz, H.: Racial differences in the use of cardiac catheterization after acute myocardial infarction. N. Engl. J. Med. **344**, 1443–1449 (2001)
9. Chen, M.K., Lowenstein, F.: The physician/population ratio as a proxy measure of the adequacy of health care. Int. J. Epidemiol. **14**, 300–303 (1985)
10. Cooper-Patrick, L., Gallo, J.J., Gonzales, J.J., Vu, H.T., Powe, N.R., Nelson, C., Ford, D.E.: Race, gender, and partnership in the patient-physician relationship. J. Am. Med. Assoc. **282**, 583–589 (1999)
11. Czypionka T, Riedel M, Röhrling G: Clinical Practice Guidelines: A Quality Instrument Comes of Age. Institute for Advanced Studies Health System Watch. http://www.ihs.ac.at/departments/fin/HealthEcon/watch/hsw06_4e.pdf (2006)
12. Dranove, D., Wehner, P.: Physician-induced demand for childbirths. J. Health Econ. **13**, 61–73 (1994)
13. Dranove, D., Satterthwaite, M.: The industrial organization of health care markets. In: Culyer, A., Newhouse, J.P. (eds.) The Handbook of Health Economics. Amsterdam, North Holland (2000)
14. Garber AM, Skinner J: Is American Health Care Uniquely Inefficient? NBER working paper 14257 (2008)
15. Getzen, T.: Health Economics and Financing, 4th edn. Wiley, Hoboken (2010)
16. Gohlke, H., Kübler, W., Mathes, P., Meinertz, T., Schuler, G., Gysan, D.B., Sauer, G.: Positionspapier zur Primärintervention kardiovaskulärer Erkrankungen. Zeitschrift für Kardiologie **92**, 522–524 (2003)
17. Greenberg, J.O., Greenberg, H.: More physicians are not the answer. Am. J. Cardiol. **99**, 1476–1478 (2007)
18. Gulich M, Zeitler P: DEGAM Leitlinie Ältere Sturzpatienten. http://www.degam.de/leitlinien/sturz_druck.pdf (2004)

19. Hagist C, Kotlikoff L. Who's Going Broke? Comparing Growth in Healthcare Costs in Ten OECD Countries. NBER working paper No. 11833 (2005)
20. Hussey, P.S., Anderson, G.F., Osborn, R., Feek, C., McLaughlin, V., Millar, J., Epstein, A.: How does the quality of care compare in five countries? Health Aff. **23**, 89–99 (2004)
21. Jencks, S.N., Cuerdon, T., Burwen, D.R., Fleming, B., Houck, P.M., Kussmaul, A.E., Nilasena, D.S., Ordin, D.L., Arday, D.R.: Quality of medical care delivered to medicare beneficiaries. J. Am. Med. Assoc. **284**, 1670–1676 (2000)
22. Jensen, M.K., Chiuve, S.E., Rimm, E.B., Dethlefsen, C., Tjønneland, A., Joensen, A.M., Overvad, K.: Obesity, behavioral lifestyle factors, and risk of acute coronary events. Circulation **117**, 3062–3069 (2008)
23. Joung, I.M.A., van der Meer, J.B.W., Mackenbach, J.P.: Marital status and health care utilization. Int. J. Epidemiol. **24**, 575–596 (1995)
24. Jürges H. Health Insurance Status and Physician-Induced Demand for Medical Services in Germany: New Evidence from Combined District and Individual Level Data. MEA working paper, pp. 119–2007 (2007)
25. Karlsson, M.: Quality incentives for GPs in a regulated market. J. Health Econ. **26**, 699–720 (2007)
26. Katz, A., Soodeen, R.A., Bogdanovic, B., De Coster, C., Chateau, D.: Can the quality of care in family practice be measured using administrative data? Health Serv. Res. **41**, 2238–2254 (2006)
27. Krakauer, H., Jacoby, I., Millman, M., Lukomnik, J.E.: Physician impact on hospital admission and on mortality rates in the medicare population. Health Serv. Res. **31**, 191–211 (1996)
28. Krämer, W.: Eine ökonometrische Untersuchung des Marktes für ambulante kassenärztliche Leistungen. Zeitschrift für die gesamte Staatswissenschaft **1**, 45–61 (1981)
29. Laditka, J.N.: Physician supply, physician diversity, and outcomes of primary health care for older persons in the United States. Health Place **10**, 231–244 (2004)
30. Laditka, J.N., Laditka, S.B., Probst, J.C.: More may be better: evidence of a negative relationship between physician supply and hospitalization for ambulatory care sensitive conditions. Health Serv. Res. **40**, 1148–1166 (2005)
31. Löwel, H., Meisinger, C.: Epidemiologie und demographische Entwicklung am Beispiel kardio-vaskulärer Erkrankungen in Deutschland. Med. Klin. **101**, 804–811 (2006)
32. McGlynn, E.A.: Six challenges in measuring the quality of health care. Health Aff. **16**, 7–21 (1997)
33. McGlynn, E.A.: There is no perfect health system. Health Aff. **23**, 100–102 (2004)
34. McGlynn, E.A., Asch, S.M., Adams, J., Keesey, J., Hicks, J., DeCristofaro, J., Kerr, E.A.: The quality of health care delivered to adults in the United States. N. Engl. J. Med. **348**, 2635–2645 (2003)
35. McGlynn, E.A., Brook, R.B.: Keeping quality on the policy agenda. Health Aff. **20**, 82–90 (2001)
36. McGuire, T.G.: Physician agency. In: Culyer, A.J., Newhouse, J.P. (eds.) Handbook of Health Economics, vol. 1. Amsterdam, North-Holland (2000)
37. McKinstry, B., Yang, S.Y.: Do patients care about the age of their general practitioner? A questionnaire survey in five practices. Br. J. Gen. Pract. **44**, 349–351 (1994)
38. Morris, S., Gravelle, H.: GP supply and obesity. J. Health Econ. **27**, 1357–1367 (2008)
39. Papke, L.E., Wooldridge, J.M.: Econometric methods for fractional response variables with an application to 401(k) plan participation rates. J. Appl. Econom. **11**, 619–632 (1996)
40. Perrin, J.M., Valvona, J.: Does increased physician supply affect quality of care? Health Aff. **5**, 63–72 (1986)
41. Pohlmeier, W., Ulrich, V.: An econometric model of the two-part decision making process in the demand for health care. J. Human Resour. **30**, 339–361 (1995)
42. Raina, P., Torrance-Rynard, V., Wong, M., Woodward, C.: Agreement between self-reported and routinely collected healthcare utilization data among seniors. Health Serv. Res. **37**, 751–774 (2002)
43. Rizzo, J.A., Zeckhauser, R.J.: Advertising and the price, quantity, and quality of primary care physician services. J. Human Resour. **27**, 381–421 (1992)
44. Scheffer, A.C., Schuurmans, M.J., van Dyck, N., van der Hooft, T., de Rooij, S.: Fear of falling: measurement strategy, prevalence, risk factors and consequences among older persons. Age Ageing **37**, 19–24 (2008)
45. Schindelheim, G., Jerrard, D., Witting, M.: Patient preference for emergency physician age and gender. Am. J. Emerg. Med. **22**, 503 (2004)
46. Shi, L., Starfield, B.: The effect of primary care physician supply and income inequality on mortality among blacks and whites in U.S. metropolitan areas. Am. J. Public Health **91**, 1246–1250 (2001)
47. Staiger, D., Stock, J.H.: Instrumental variables regression with weak instruments. Econometrica **65**, 557–586 (1997)
48. Stevens, J.A., Mack, K.A., Paulozzi, L.J., Ballesteros, M.F.: Self-reported falls and fall-related injuries among persons aged $\geq$ 65 years—United States, 2006. J. Am. Med. Assoc. **299**, 225–229 (2008)
49. Vellas, B.J., Wayne, S.J., Romer, L.J., Baumgarner, R.N., Garry, P.J.: Fear of falling and restriction of mobility in elderly fallers. Age Ageing **26**, 189–193 (1997)
50. Wennberg JE, Fisher ES, Skinner JS: Geography and the debate over medicare reform. Health Affairs Suppl. Web Exclusives. W96–114 (2002)