

Danny Wende
Thomas Kopetsch
Wolfram F. Richter

Planning Health Care Capacities with a Gravity Equation

Imprint

Ruhr Economic Papers

Published by

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

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

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

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

Editors

Prof. Dr. Thomas K. Bauer

RUB, Department of Economics, Empirical Economics
Phone: +49 (0) 234/3 22 83 41, e-mail: thomas.bauer@rub.de

Prof. Dr. Wolfgang Leininger

Technische Universität Dortmund, Department of Economic and Social Sciences
Economics – Microeconomics
Phone: +49 (0) 231/7 55-3297, e-mail: W.Leininger@tu-dortmund.de

Prof. Dr. Volker Clausen

University of Duisburg-Essen, Department of Economics
International Economics
Phone: +49 (0) 201/1 83-3655, e-mail: vclausen@vwl.uni-due.de

Prof. Dr. Ronald Bachmann, Prof. Dr. Manuel Frondel, Prof. Dr. Torsten Schmidt,
Prof. Dr. Ansgar Wübker

RWI, Phone: +49 (0) 201/81 49-213, e-mail: presse@rwi-essen.de

Editorial Office

Sabine Weiler

RWI, Phone: +49 (0) 201/81 49-213, e-mail: sabine.weiler@rwi-essen.de

Ruhr Economic Papers #888

Responsible Editor: Wolfgang Leininger

All rights reserved. Essen, Germany, 2020

ISSN 1864-4872 (online) – ISBN 978-3-96973-027-0

The working papers published in the series constitute work in progress circulated to stimulate discussion and critical comments. Views expressed represent exclusively the authors' own opinions and do not necessarily reflect those of the editors.

Ruhr Economic Papers #888

Danny Wende

Thomas Kopetsch

Wolfram F. Richter

**Planning Health Care Capacities
with a Gravity Equation**

Bibliografische Informationen der Deutschen Nationalbibliothek

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

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

<http://dx.doi.org/10.4419/969730270>

ISSN 1864-4872 (online)

ISBN 978-3-96973-027-0

Danny Wende,¹ Thomas Kopetsch² and Wolfram F. Richter³

Planning Health Care Capacities with a Gravity Equation

Abstract

The planning of health care capacities is in practice constrained by sectoral and regional boundaries and it remains difficult to ensure an adequate and even access to health care. Moreover, standard planning approaches lack the choice-theoretic grounding necessary for making reliable statements about the demand for health care. This paper presents a model based on the idea of gravity in supply and demand linkages and designed to overcome such shortcomings. Empirical estimation equations are derived for the size of catchment areas, the spatial access to health care and the demand for specialist treatment. The floating catchment area (FCA) method commonly used to measure access to care is shown to be a special, yet often misleading case. This is demonstrated by the example of Germany, where rural areas are shown to suffer from access deficits.

JEL-Code: H72, I18, D58, C3

Keywords: Health care capacity planning; choice-theoretic foundation; gravity equation; spatial multiple equation model

December 2020

¹ Danny Wende, TU Dresden, c/o Chair for Quantitative Methods; ² Thomas Kopetsch, Formerly: National Association of Statutory Health Insurance Physicians, Berlin; ³ Wolfram F. Richter, TU Dortmund, IWH Halle, CESifo Munich, and IZA Bonn. All correspondence to: Danny Wende, Burgkstr. 31, 01159 Dresden, Germany. Email: danny.wende@bifg.de

1 Introduction

All well-developed health systems use capacity planning to ensure adequate access to care and a balanced allocation of resources. Health care capacity planning in its true sense is a forecast of a population's medical needs by geographical region. It aims to provide a basis for policy decisions that avoid overuse, underuse and misuse of medical resources (Ono et al., 2013).

The medical needs of the population are expressed in the demand from patients, which is in part exogenous but also depends on factors that are subjective and susceptible to social influence. Capacity planning uses appropriate indicators of population characteristics, such as age and gender distribution, to estimate capacity needs (Ono et al., 2013). Yet although in reality health care is characterised by a transfer of services between regions and sectors, especially between urban and rural areas and between the ambulatory and hospital care sectors, capacity planning is almost without exception area- and sector-specific (Ono et al., 2013). In addition, most planning models take only rudimentary account of demand structures with cross-border effects both within and between sectors. Czihal et al. (2012) have shown that there are regions of Germany which, as net importers of health services, meet only 39% of their local demand. By contrast, net exporters deliver up to 473% of services used locally. All large European countries face similar challenges in their capacity planning.

The present paper draws on models developed over the last fifty years to explain interregional and intersectoral trade in goods and services. These spatial-economic models, known as gravity models, can predict such trade very accurately (Anderson, 2011). Although they are already in use in health economics, for example for measuring access to health care and predicting patient flows, a truly convincing connection between capacity planning and gravity theory has yet to be made. Despite their undeniable advantages, gravity models are not used in capacity planning, while studies which use them to measure access to health resources often rely on empirically poorly validated parameter assumptions, such as the size of catchment areas and cross-sectoral effects (see following section).

The objective of this paper is to make gravity models fit for use in capacity planning. With this in mind, the paper is divided into two sections. The first presents a full equilibrium model which is grounded in choice theory and features the idea of gravity in supply and demand linkages. The model combines the approach developed by Bikker and de Vos (1992) with the trade model developed by Anderson and van Wincoop (2003). In the second section, we take Germany as an example to show that our model is well-suited to practical capacity planning.

Importantly, we uncover imbalances in the access to health care, showing rural and deprived regions to be undersupplied with health care resources while oversupply is a characteristic of urban areas.

2 Related literature

Gravity models are derived from the law of gravity in classical physics, which states that two bodies attract each other and that the strength of the attraction depends on the masses of the bodies and their physical distance (Anderson, 2011). This basic idea is taken up by economic gravity models. The first theoretical modelling in the context of international trade was performed by Anderson (1979),¹ who derived the gravity equation by assuming regionally distributed consumers and producers of commodities. The consumers maximize utility at constant elasticity of substitution (CES), the producers maximize profit, and interregional trade is costly. It follows from the model that, after controlling for size differences, trade between two regions depends on the ratio of bilateral trade costs to the average of all trade costs. Variants of the model in more recent publications supplement the assumption of CES preferences with an assumption of monopolistic competition in order to endogenize the specialization of producers. An early example of this is Bergstrand (1985).

The health economics literature often refers to the first concept of gravity, which appears in Huff's (1964) description of catchment areas. Lowe and Sen (1996) use a gravity equation to predict patient flows and catchment areas. The commonest approach to measuring access to health care tries to capture the idea of gravity in supply-demand relationships by building on the floating catchment area (FCA) method² (Delamater, 2013). However, this approach suffers from a strong supply bias, with the demand side represented only by functional *ad-hoc* specifications. As a consequence, estimates of model parameters are often determined anecdotally. For example, catchment areas are simply set at 30 to 60 minutes or at corresponding distances (Delamater, 2013), or are replaced by proxy variables (Matthews et al., 2019). The present paper seeks progress by deriving a gravity equation that is choice-theoretically grounded and shown to be well-suited to planning sector- and region-specific health care capacities. The model presented combines the approach to measuring access to

¹ It should be noted, however, that Leamer and Stern (1970) had already provided the trade gravity equation by intuitive reasoning.

² Variants of the FCA method are the two-step (2S-), the extended two-step (E2S-), and the kernel density two-step (KD2S-) FCA methods. They mainly differ from the base version in spatial weighting (Delamater, 2013). The integrated (i-) and the three-step (3S-) FCA methods will be discussed in Section 3.5.

health care taken by Bikker and de Vos (1992) with a trade model developed by Anderson and van Wincoop (2003), building on Bergstrand (1985).

3 Modelling supply and demand for health care

Health care capacities are assumed to differ by the place of service delivery, $j \in \{1, \dots, \#j\}$, and the type of medical speciality, $s \in \{1, \dots, \#s\}$. They are demanded and utilised by individuals as patients. An individual's place of residence is not necessarily the place where health services are delivered. On the contrary, it is assumed that individuals live in various different places, $i \in \{1, \dots, \#i\}$, and travel to various locations, j , for treatment. Let the spatial distance between the place of residence i and the place of service delivery j be denoted by d_{ij} .

3.1 The demand for health care

An individual with residence at i expects to derive utility from health-neutral consumption c_i and a demanded bundle of health services. The sub-utility derived from the bundle is denoted by V . By assumption, the elasticity of substitution between health services and health-neutral consumption is one (Cobb-Douglas) and the partial elasticity of health-neutral consumption is $1 - \mu_i$. The parameter μ_i captures the time individuals with residence at i spend on the consumption of health services and can be interpreted as the morbidity rate. The sub-utility V is a function of the quantities of contacts, n_{ijs} , that the patient has with the service providers of speciality s at location j . The specification of utility is assumed to be the same for all individuals:

$$U(c_i, n_{i.}) \equiv c_i^{1-\mu_i} \cdot V^{\mu_i} \quad \text{with}$$

$$V(n_{i.}) \equiv (\sum_s \sum_j \vartheta_s \cdot n_{ijs}^{(\sigma-1)/\sigma})^{\sigma/(\sigma-1)} \quad \text{and} \quad n_{i.} = (n_{ijs})_{js} \equiv (n_{i11}, \dots, n_{ijs}, \dots). \quad (1)$$

The sub-utility V expresses an appreciation of diversity and thus an individual's willingness to substitute between locations and specialities. The parameter σ measures the elasticity of substitution, which is assumed to be constant (CES function). A low value for σ indicates that health services are poorly substitutable in demand, whereas a high value indicates easy substitution. The preferential weight given to different specialities is captured by ϑ_s . The higher this parameter is, the more weight is given to the corresponding speciality.

Individuals have a time budget which is used for labour supply, l_i , and doctor contacts, n_{ijs} . The gross time of treatment, t_{ijs} , includes both the time the patient needs to visit the doctor's location, $f_{ijs} \equiv f_s(d_{ij})$, and the time, τ_{js} , needed at location j for treatment in speciality s . For model-related reasons, the two time components are assumed to be multiplicatively connected, $t_{ijs} \equiv \tau_{js} \cdot f_{ijs}$. Unlike the treatment times, the distances are assumed to be unalterable. Hence f_{ijs} is exogenous to the model whereas τ_{js} is an endogenous variable and the equilibrium result of supply and demand for treatment. The price of consumption is normalized to one and the real wage rate is w .

$$c_i = w \cdot l_i \quad (\text{consumption budget}) \quad (2)$$

$$\sum_{js} n_{ijs} \cdot \tau_{js} \cdot f_{ijs} = y - l_i \quad (\text{time budget}). \quad (3)$$

The parameter y can be read as “year”. By maximizing the utility function (1) subject to the constraints (2) and (3) we obtain the supply of labour, l_i , as a fraction of the year, the demand for doctor contacts, $n_{ijs} = n_{ijs}(\tau.)$, as a function of the treatment-time profile, $\tau. \equiv (\tau_{js})_{js}$, and the demand for health-neutral consumption, $c_i = c_i(w)$, as a function of the real wage rate, w :

$$l_i = (1 - \mu_i)y, \quad n_{ijs}(\tau.) = \mu_i y \cdot \vartheta_s^\sigma \cdot \frac{\tau_{js}^{-\sigma} \cdot f_{ijs}^{-\sigma}}{T_i(\tau.)}, \quad c_i(w) = w(1 - \mu_i)y. \quad (4)$$

According to eq. (3), $\mu_i y = y - l_i$ is the time spent on health care. The denominator,

$$T_i(\tau.) \equiv \sum_{js} \vartheta_s^\sigma \cdot \tau_{js}^{1-\sigma} \cdot f_{ijs}^{1-\sigma}, \quad (5)$$

can be interpreted as a (preference-weighted) index of gross treatment times. Its role is to ensure that the individual keeps to the time budget. This is so for the following reason. Suppose that treatment times and distances are small. This will have an increasing effect on the demand for doctor contacts. A large value of T_i is then required to bring the demand for care into line with the time budget. It is worth noting that T_i is independent of $\tau.$ and constant in i if $\sigma = 1$ (Cobb-Douglas). As spatial distances are unalterable by assumption, we suppress the dependence of T_i and n_{ijs} on f_{ijs} .

3.2 The supply of medical treatment and general equilibrium conditions

In many countries, the remuneration of health care services is strictly regulated. In such circumstances, doctors' scope of action is limited to the amount of labour they choose to

supply. In its simplest form, the amount L_{js} chosen by a doctor practicing in speciality s at location j is merely a function of the doctor's real fee, h_{js} . Total labour supply, S_{js} , amounts to individual labour supply, $L_{js}(h_{js})$, multiplied by the number of doctors, A_{js} .

Patients demand doctor contacts which consume time. Let I_i be the number of individuals with residence i and $x_{ijs} = x_{ijs}(\tau.) \equiv I_i \cdot n_{ijs}(\tau.) \cdot \tau_{js}$ the time demanded by those individuals for treatment in speciality s at location j . In equilibrium, the demand is met by supply,

$$\sum_i x_{ijs} = \sum_i I_i \cdot n_{ijs} \cdot \tau_{js} = A_{js} \cdot L_{js} \equiv S_{js} \quad \text{for all } js. \quad (6)$$

Consumer goods must be produced and we assume a form of production in which consumer goods are produced with labour alone at constant returns to scale. In equilibrium, the supply price is determined by the cost of labour. Health services are financed by a payroll tax at rate b . Thus $1 - b = w$ must apply. The balancing of revenues and expenditures in the health budget requires

$$b \cdot \sum_i I_i \cdot l_i = \sum_{js} A_{js} \cdot h_{js} \cdot L_{js}. \quad (7)$$

The right-hand side of eq. (7) represents doctors' demand for consumption goods. By Walras' Law this demand is equal to $\sum_i I_i \cdot (l_i - c_i)$, which can be interpreted as the individuals' excess supply of consumption goods. The payroll tax rate, b , balances the budget.

3.3 A gravity-theoretic representation of the demand for medical treatment

The spatial distribution of the demand for medical treatment is given by x_{ijs} . In order to investigate this distribution more closely, we introduce two auxiliary terms α_{js} and β_{is} . As a first step, we set

$$\beta_{is} = \beta_{is}(\tau.) \equiv \sum_j \tau_{js}^{1-\sigma} \cdot f_{ijs}^{-\sigma}. \quad (8)$$

β_{is} can be interpreted as an index of access to care (Delamater, 2013). To see this, it needs to be noted that treatment times, τ_{js} , are constant in i . If they are also constant in js , β_{is} can only assume a large value, because some values of $f_{ijs}^{-\sigma}$ are large. This means that some distances f_{ijs} to doctors must be small, thus indicating easy access to care. If treatment times are not constant, the distances enter the definition of access in weighted form. This weighting, however, depends on the substitutability of health services. If substitutability is low, $\sigma < 1$, and the treatment time long, a short distance is highly weighted. In the alternative case

characterized by high substitutability, a short distance is highly weighted only if the treatment time is short. A notable case is where $\sigma = 1$. In this case of Cobb-Douglas preferences, β_{is} is independent of treatment times even if they are not constant.

The spatially aggregated demand of individuals for treatment in specialization s at location i can be written as

$$D_{is} = D_{is}(\tau.) \equiv \sum_j x_{ijs} = \vartheta_s^\sigma \cdot \tilde{I}_i \cdot \beta_{is} \quad \text{with} \quad \tilde{I}_i = \tilde{I}_i(\tau.) \equiv \frac{\mu_i y}{T_i(\tau.)} I_i. \quad (9)$$

Demand at location i thus increases proportionally with the product of \tilde{I}_i and the index of access to care, β_{is} . The structure of the equation is as in Bikker and de Vos (1992) except for the factor $\mu_i y / T_i$, which can be interpreted as a population adjustment with no direct analogue in their approach. The adjustment is made with regard to the morbidity rate, μ_i , and the index of gross treatment times, T_i . Let us call \tilde{I}_i the number of individuals at location i adjusted for morbidity and distance. In the literature on FCA, $D_{is} / \beta_{is} = \vartheta_s^\sigma \cdot \tilde{I}_i$ is interpreted as (access-independent) demand potential (Delamater, 2013). Combining equations (4), (8), and (9), we obtain

$$\phi_{ijs} = \phi_{ijs}(\tau.) \equiv \frac{x_{ijs}}{D_{is}} = \frac{\tau_{js}^{1-\sigma} \cdot f_{ijs}^{-\sigma}}{\beta_{is}}. \quad (10)$$

In the literature, ϕ_{ijs} is interpreted as a cross-border supply ratio (Czihal et al., 2012). It measures the proportion of i 's aggregate demand for treatment in s met by supply from j .

In a second step, we set

$$\alpha_{js} = \alpha_{js}(\tau.) \equiv \vartheta_s^\sigma \cdot \sum_i \tilde{I}_i(\tau.) \cdot f_{ijs}^{-\sigma} \quad (11)$$

so that the supply of treatment time in speciality s at location j can be written as

$$S_{js} = \sum_i x_{ijs} = \vartheta_s^\sigma \cdot \tau_{js}^{1-\sigma} \cdot \sum_i \tilde{I}_i \cdot f_{ijs}^{-\sigma} = \tau_{js}^{1-\sigma} \cdot \alpha_{js}. \quad (12)$$

From this equation it can be seen that α_{js} measures the number of treatments carried out in speciality s at location j . In the literature on FCA, α_{js} is interpreted as the population's care potential and S_{js} / α_{js} as the supply ratio (Delamater, 2013).

From equations (10) and (12) we obtain

$$x_{ijs} = \frac{D_{is}}{\beta_{is}} \cdot \tau_{js}^{1-\sigma} \cdot f_{ijs}^{-\sigma} = \left(\frac{D_{is}}{\beta_{is}} \right) \cdot \left(\frac{S_{js}}{\alpha_{js}} \right) \cdot f_{ijs}^{-\sigma}. \quad (13)$$

The demand of individuals with residence i for treatment in speciality s at location j , x_{ijs} , is thus proportional to the "masses" given by the demand potential and the supply ratio. Furthermore, it is proportional to an indicator of distance. Eq. (13) thus features a relationship which can be given a gravity-theoretic interpretation.

Following Bikker and de Vos (1992) we use eq. (13) to express the index of access to care, β_{is} , as a function of $D_{is}, x_{ijs}, f_{ijs}^{-\sigma}$. In doing so we take advantage of the fact that the indices are only uniquely determined up to a joint scalar. Note that we are able to replace β_{is} and α_{js} in eq. (13) with $k_s \cdot \beta_{is}$ and α_{js}/k_s , respectively, without changing anything else. Hence, we can set k_s at a value satisfying the constraint of $\sqrt[\#i]{\prod_i \beta_{is}} = 1$. By calculating a geometric mean of eq. (13) with respect to i we obtain the equation $\frac{s_{js}}{\alpha_{js}} = \sqrt[\#i]{\prod_i \frac{x_{ijs} \cdot f_{ijs}^{\sigma}}{D_{is}}}$. By inserting this equation into eq. (13) and calculating a further geometric mean with respect to j we obtain a formulaic representation of β_{is} :

$$\beta_{is} = \sqrt[\#j]{\prod_j \frac{\sqrt[\#i]{\prod_i x_{ijs} \cdot f_{ijs}^{\sigma}}}{x_{ijs} \cdot f_{ijs}^{\sigma} \sqrt[\#i]{\prod_i D_{is}}}} \cdot D_{is}. \quad (14)$$

This formula will later be used for empirical measurement. Setting $k \equiv [\prod_j (\prod_i x_{ijs} \cdot f_{ijs}^{\sigma} / D_{is})^{\frac{1}{\#i}}]^{\frac{1}{\#j}}$ the index can be rewritten as

$$\beta_{is} = k \cdot \frac{D_{is}}{\sqrt[\#j]{\prod_j x_{ijs} \cdot f_{ijs}^{\sigma}}} = k \cdot [(\prod_j \phi_{ijs})^{\frac{1}{\#j}} \cdot (\prod_j f_{ijs})^{\frac{\sigma}{\#j}}]^{-1}. \quad (15)$$

According to this formula, β_{is} is inversely proportional to the product of the geometric means of cross-border supply ratios and distances. This means that the access to care is negatively affected by high cross-border supply ratios and long distances to service providers. Furthermore, each geometric mean can be substituted by the other at constant elasticity.

3.4 Planning even access to care

Health equity is of paramount importance in any conception of social justice (Sen, 2002).³ This is recognized in most countries and is the reason why health authorities pay the greatest

³ For a nuanced discussion of equity in health care, see Fleurbaey and Schokkaert (2009). These authors distinguish between ethically legitimate and illegitimate inequities, the latter being characterized by the fact that they are beyond individual control. Often, the causes work to the disadvantage of the socially deprived segments of the population (Vogt, 2016).

attention to health care capacity planning and resource allocation (Ono et al., 2013). Taking equitable or evenly distributed access to care as a norm, a mandate can be derived to plan health care capacities in such a way that the targeted indices of access, $\bar{\beta}_{is}$, are independent of place of residence and thus assume equal values,

$$\bar{\beta}_s = \bar{\beta}_{is} = \sum_j \bar{\tau}_{js}^{1-\sigma} \cdot f_{ijs}^{-\sigma} \quad \text{for all } is. \quad (16)$$

The model suggests that this equality can be achieved by adjusting treatment times, $\bar{\tau}_{js}$, and/or travelling times, f_{ijs} . In principle, one might consider shortening physical distances, d_{ij} , for example by means of patient transport services or improved public transport connections. Even psychological distances, which are rooted in a lack of information, can be reduced by appropriate measures. In the short term, however, physical distances remain immutable and the only way of varying $\bar{\beta}_{is}$ is by assuming $\sigma \neq 1$ and varying $\bar{\tau}_{js}$. This entails the exclusion of Cobb-Douglas preferences in what follows. We next prove that an even access to care requires an allocation of resources such that the aggregate supply of treatment time is allocated to the places of residence in proportion to a distance-adjusted number of individuals (Proposition 1). Stating and proving this in precise terms requires some preparatory considerations.

Eq. (16) defines a system of equations which under conditions of regularity can be used to determine the matrix $\bar{\tau}^{1-\sigma} \equiv (\bar{\tau}_{js}^{1-\sigma})_{js} = (\bar{\tau}_{js}^{1-\sigma}(\bar{\beta}_s))_{js}$ as a function of $\bar{\beta}_s$.⁴ By the system's linear structure, $\bar{\tau}_{js}^{1-\sigma}(\bar{\beta}_s) = \bar{\beta}_s \cdot \bar{\tau}_{js}^{1-\sigma}(1)$ must hold. Note that the legal obligation to provide even access to health care leaves the question about the level of access open. We assume that the answer – the choice of $\bar{\beta}_s$ – is a political one. Associated with $\bar{\beta} = (\bar{\beta}_s)_s$, there is the index of gross treatment times,

$$\bar{T}_i(\bar{\beta}) \equiv T_i(\bar{\tau}(\bar{\beta})) = \sum_{js} \vartheta_s^\sigma \cdot \bar{\tau}_{js}^{1-\sigma}(\bar{\beta}_s) \cdot f_{ijs}^{-\sigma} = \sum_{js} \vartheta_s^\sigma \cdot \bar{\beta}_s \cdot \bar{\tau}_{js}^{1-\sigma}(1) \cdot f_{ijs}^{-\sigma}. \quad (17)$$

Let $\bar{I}_i \equiv \tilde{I}_i(\bar{\beta}) = \frac{\mu_i \gamma}{\bar{T}_i(\bar{\beta})} I_i$ denote the morbidity- and distance-adjusted number of individuals associated with $\bar{\beta}$.

Any planned adjustment of net treatment times requires balancing supply and demand. Of the two sides, only the supply, \bar{S}_s , can be the direct object of planning. By contrast, demand at location i , \bar{D}_{is} , will react endogenously to the changes in treatment times. The variables that

⁴ Solvability requires invertible matrices $(f_{ijs}^{-\sigma})_{ij}$ for all s and *a fortiori* equality of $\#j$ and $\#i$ (Gentle, 2007, p. 211).

allow planners to control supply are the number and location of service providers, \bar{A}_{js} , and doctors' real fees, \bar{h}_{js} . Adjustment of the latter will change the supply of doctors' labour. The choice of \bar{A}_{js} and \bar{h}_{js} must be such that the demand for treatment and its supply are balanced,

$$\sum_i x_{ijs}(\bar{\tau}.) = \bar{S}_{js} = S_{js}(\bar{\tau}., (\bar{\beta}..)) = \bar{A}_{js} \cdot L_{js}(\bar{h}_{js}) \quad \text{for all } js. \quad (18)$$

Changes in the number of service providers and their fees have an effect on expenditures. Budget balancing requires a corresponding adjustment of revenues (see eq. (7)). The variable with which revenues can be adjusted is the payroll tax rate, \bar{b} .

Proposition 1: An even access to health care, $\bar{\beta}_{is} = \bar{\beta}_s$, requires an allocation of resources such that the aggregate supply of treatment time, \bar{S}_s , is allocated to the places of residence in proportion to the morbidity- and distance-adjusted number of individuals,

$$\bar{D}_{is} = \frac{\bar{I}_i}{\sum_{i'} \bar{I}_{i'}} \cdot \bar{S}_s. \quad (19)$$

The proof is straightforward. From the equations (12) and (16) we obtain

$$\bar{S}_s = \sum_j \bar{S}_{js} = \vartheta_s^\sigma \cdot \sum_i \bar{I}_i \cdot \sum_j \bar{\tau}_{js}^{1-\sigma} \cdot f_{ijs}^{-\sigma} = \vartheta_s^\sigma \cdot \sum_{i'} \bar{I}_{i'} \cdot \bar{\beta}_s \quad (20)$$

and from equations (9) and (16)

$$\bar{D}_{is} = \vartheta_s^\sigma \cdot \bar{I}_i \cdot \bar{\beta}_s. \quad (21)$$

Eq. (19) follows by dividing eq. (21) by eq. (20). \square

If the preference weights, ϑ_s , are known, equations (19) - (21) can be used for guiding resource planning. Just note that eq. (20) implies

$$\frac{\bar{S}_s}{\bar{S}_1} = \frac{\vartheta_s^\sigma \bar{\beta}_s}{\vartheta_1^\sigma \bar{\beta}_1} \quad \text{for } s \neq 1, \quad (22)$$

i.e. planned relative supplies, \bar{S}_s/\bar{S}_1 , need to equate the product of planned relative accesses to care, $\bar{\beta}_s/\bar{\beta}_1$, and of the relative weight, $\vartheta_s^\sigma/\vartheta_1^\sigma$, given by individuals to specialities. After substituting $\vartheta_s^\sigma \bar{\beta}_s$ by $\bar{S}_s \cdot \vartheta_1^\sigma \bar{\beta}_1/\bar{S}_1$ in equations (20) and (17) and reducing, we obtain the equation

$$1 = \sum_i \frac{\mu_i \gamma_i I_i}{\sum_{js} \bar{S}_s \cdot \bar{\tau}_{js}^{1-\sigma} (1) \cdot f_{ijs}^{1-\sigma}}. \quad (23)$$

This tells us that the vector of planned supplies, $(\bar{S}_1, \dots, \bar{S}_s, \dots)$, must satisfy a constraint determined by the regional distributions of population, mortality, and distance, I_i, μ_i, f_{ijs} , respectively. This applies in both the static and dynamic senses. If, for example, population and mortality change as a result of demographic change, planned supplies must be adapted accordingly.

3.5 A comparison of the gravity model with competing approaches

As mentioned above, the idea of measuring access by deriving the gravity equation (13) from a supply- demand relationship goes back to Bikker and de Vos (1992). The difference between their approach and that of the present authors lies in the derivation of supply and demand for health care. Bikker and de Vos postulate the existence of such functions. They focus on a particular medical speciality and supply and demand are analysed in partial equilibrium. The present approach has a choice-theoretic grounding. Individuals must divide their time budget between different specialities, and demand and supply are analysed in full equilibrium. The difference can be seen in the index of gross treatment times, T_i , which has no analogue in the approach taken by Bikker and de Vos. This is of no great relevance when simply measuring access to care, as in the next section, but takes on greater significance when planning the allocation of resources. If it is necessary to determine how resources are best allocated to competing specialities, the restrictions on individuals' time and their willingness to substitute between specialities must be modelled consistently.

The dominant approach to measuring access to care in the literature builds on the FCA method. In its base version, the index of access, $\beta_{i,FCA}$, is defined in a two-step procedure. In the first step, the population is aggregated to a demand potential using a geographical weighting scheme, $g(d_{ij})$. In the second, the ratio of supply (often approximated by the number of service providers) to the demand potential is calculated using the geographical weighting scheme to create the index (Delamater, 2013). If we suppress s and equate $g(d_{ij})$ with $f_{ij}^{-\sigma}$ the index can be written as:

$$\beta_{i,FCA} = \sum_j \frac{S_j g(d_{ij})}{\sum_{i'} I_{i'} g(d_{i'j})} = \sum_j \frac{S_j f_{ij}^{-\sigma}}{\sum_{i'} I_{i'} f_{i'j}^{-\sigma}}. \quad (24)$$

A structurally similar formula can be derived from the gravity model by relying on equations (8), (11), and (12):

$$\beta_{i,GM} = \sum_j \tau_j^{1-\sigma} \cdot f_{ij}^{-\sigma} = \sum_j \frac{S_j}{\alpha_j} \cdot f_{ij}^{-\sigma} = \vartheta^{-\sigma} \cdot \sum_j \frac{S_j \cdot f_{ij}^{-\sigma}}{\sum_{i'} \tilde{I}_{i'} \cdot f_{i'j}^{-\sigma}} \quad (25)$$

The primary difference between the definitions is that $\beta_{i,FCA}$ is defined as a function of I_i whereas $\beta_{i,GM}$ is defined as a function of $\tilde{I}_i = \frac{\mu_i y}{T_i} I_i$. This difference does not matter if the adjustment factor, $\mu_i y / T_i$, is constant in i . This is clearly the case if both μ_i and T_i are constant in i . Constancy of T_i is ensured if individuals' preferences are of the Cobb-Douglas type. However, assuming $\sigma = 1$ and $\mu_i = \text{constant}$, the equations (12) and (11) imply $S_j = \alpha_j = k \cdot \sum_i I_i f_{ij}^{-1}$. The indices, $\beta_{i,FCA}$ and $\beta_{i,GM}$, are then equal up to a certain constant, as can easily be understood by comparing the equations (24) and (25). The FCA method and the gravity model should therefore produce identical geographic maps of access to care.

Proposition 2: The FCA method can be derived from the gravity model if the need to adjust the population for morbidity and distance is ignored. In this case, the demand potential is proportional to the population size, $D_i / \beta_i = k \cdot I_i$.

In fact, the geographic maps shown for $\beta_{i,FCA}$ and $\beta_{i,GM}$ in the next section differ considerably. From a planning perspective, therefore, a choice between the competing approaches to measuring access to care cannot be avoided. The difference between the approaches can be described as follows. The FCA method is primarily supply oriented, with the demand potential being only rudimentarily captured by the size and regional distribution of the population. Access is improved if nearby service capacities are increased. The orientation of the gravity model is better balanced between supply and demand. In particular, the effect of morbidity on demand for care is explicitly modelled. Access to care is improved if cross-border supply ratios are reduced, for example, by meeting a larger share of local demand from nearby supply.

Below, both methods of measuring access to care are empirically illustrated. It must be pointed out, however, that in the literature the FCA method is no longer regarded as state-of-the-art. Wan et al. (2012), for example, find it likely that it overestimates the demand for health care. Criticism of this kind has led to efforts to develop more refined but still supply-oriented indices of access to care. Examples that deserve special mention are the three steps

(3S-) and the integrated (i-) FCA methods proposed by Wan et al. (2012) and Luo (2014), respectively:

$$\beta_{i,3SFCA} = \sum_j \frac{S_j \cdot g(d_{ij}) \cdot G_{3SFCA}(d_{ij})}{\sum_{i'} I_{i'} \cdot g(d_{i'j}) \cdot G_{3SFCA}(d_{i'j})} \text{ with } G_{3SFCA}(d_{ij}) = \frac{g(d_{ij})}{\sum_{j': d_{ij'} < \bar{d}} g(d_{ij'})} \quad (26)$$

$$\beta_{i,iFCA} = \sum_j \frac{S_j \cdot g(d_{ij})}{\sum_{i'} I_{i'} \cdot g(d_{i'j}) \cdot G_{iFCA}(S_j, d_{i'j})} \text{ with } G_{iFCA}(S_j, d_{ij}) = \frac{S_j \cdot g(d_{ij})}{\sum_{j': d_{ij'} < \bar{d}} S_{j'} \cdot g(d_{ij'})} \quad (27)$$

where \bar{d} denotes a critical distance. The obvious difference to the FCA method comes from the factor G_{*FCA} , which is added to capture “the availability of opportunities as moderated by distance” (Delamater, 2013). More precisely, the 3SFCA method incorporates the assumption that the demand for medical treatment is proportional to weighted distances. By contrast, the iFCA method aims at improving the standard FCA method by incorporating the model developed by Huff (1964), which allows for population selections of healthcare services within a catchment area (Luo, 2014). The Huff model assumes high demand if nearby supply is strong. After running simulations, Paez et al. (2019) conclude that access to care is best measured by the iFCA method. However, all attempts to measure access to care based on the FCA method are open to the same criticism, namely that they are supply-biased. The demand side is only taken into account through functional refinements largely justified by ad-hoc considerations. In contrast to this paper’s gravity model, they suffer from the lack of a firm choice-theoretic grounding.

4 Measuring access to health care empirically

4.1 Sector-specific capacity planning in Germany

Germany lends itself to a study of health care capacity planning because the concept of “equal living standards” across the whole country has constitutional status (Article 72 of the Basic Law of Germany). Access to health care must therefore be made available to all citizens equally and evenly (Sundmacher et al., 2018). This right is expressed in a free choice of doctors and in the almost complete assumption of costs for all services offered by the statutory health insurance (SHI) system. It has to be noted, however, that there is a strict separation between ambulatory and hospital care in Germany. Ambulatory capacity planning is carried out nationally, but at a small-scale level, on the basis of the so-called Capacity Planning Guideline (Bedarfsplanungs-Richtlinie) drawn up by the Federal Joint Committee (Gemeinsamer Bundesausschuss), comprised of representatives of both the statutory health

insurers and service providers. The planning structure is supply-oriented. The reference values against which current doctor-density ratios are measured are the ratios of doctors to population in 23 different medical specialties determined on a certain date in the past. They constitute the policy target for the number of ambulatory physicians in a full-time position to be financed by the SHI system.

In contrast, hospital capacities are planned at the level of Germany's sixteen federal states. The need for hospital beds is estimated on the basis of the Hill-Burton formula, with the target number of beds being updated from actual figures. This method is known to have the effect that the greatest capacity need is assessed where the existing capacity is already high.

The strict separation between ambulatory and hospital care is a characteristic feature of the German health care system and means that specialist capacities are available in both sectors. In practice, there is much overlap that is only partially taken into account in planning and ultimately leads to efficiency losses (Kopetsch, 2007b). The separation of ambulatory and hospital capacity planning causes particular problems for rural areas, as hospital planning aims at exploiting economies of scale and thus gives priority to urban areas.

4.2 German health care data

For our empirical analysis we use data for the ambulatory sector provided by the National Association of Statutory Health Insurance Physicians (Kassenärztliche Bundesvereinigung, KBV). We use two primary data sources. The first is the Federal Register of Physicians, a directory of all doctors accredited to the SHI. Secondly, we use the KBV's billing data, which includes information on the patient treated, the diagnosis, the services provided, and the doctor's specialization. The geographical information for both doctors and patients is given by postal codes. For hospital care we use the hospitals' quality reports, which contain the number of cases per hospital with a postal code reference. The values used in hospital billing are known as diagnosis-related groups (DRGs) and from an analysis of invoiced DRGs specially prepared by the Federal Statistical Office (Destatis) we also have access to hospital cases per patient at the county level. We determine the minimum travel times between patients' homes and physicians' locations by using OpenStreetMap road data.

The quantity of medical services provided is measured by the number of treated cases⁵ billed by general practitioners, self-employed specialists and hospitals (see Table 1). The time required per case is 9.9 minutes in ambulatory care and 7.2 days in hospital care. The variable x_{ijs} is determined by the time required per case and the number of cases performed at location j by a supplier of type s for patients residing in i . Aggregate supply is $S_{js} = \sum_i x_{ijs}$, and aggregate demand is $D_{is} = \sum_j x_{ijs}$.

Table 1: Billed ambulatory and hospital cases, number of office-based doctors and of hospital beds

Specialization	Doctors / Beds 2014	Doctors / Beds 2015	Cases 2014	Cases 2015
General Practitioners	51,808	51,628	205,601,758	207,063,055
Ophthalmologists	5,277	5,279	28,325,853	27,990,509
Surgeons	3,890	3,880	10,693,797	10,255,080
Gynaecologists	9,757	9,748	43,653,927	43,013,618
ENT-specialists	3,886	3,910	18,458,521	18,627,133
Dermatologists	3,220	3,209	20,754,040	20,486,739
Paediatricians	5,627	5,613	23,067,755	23,266,578
Neurologists	4,899	4,827	10,516,440	10,329,799
Orthopaedic specialists	5,317	5,370	21,292,427	20,988,025
Psychotherapists	22,804	23,264	4,671,210	4,908,858
Urologists	2,650	2,644	12,283,758	12,346,443
Hospital beds	638,064	636,407	17,877,555	18,304,621

Note: The number of doctors is given in full-time equivalents

4.3 Estimating distances to care

Our measurement of access to care relies on eq. (14). This requires an estimation of the distances to doctors' locations, $f_{ijs} \equiv f_s(d_{ij})$. We use a parameterization based on a Gaussian variogram model (Chilès and Delfiner, 1999):

$$(f_{ijs}(d_{ij}))^{-\sigma} = \left(\exp\left(\frac{d_{ij}^2}{\mu\phi_s^2}\right)\right)^{-\sigma} = \exp(\theta_s \cdot d_{ij}^2) . \quad (28)$$

The parameter $\theta_s \equiv -\frac{\sigma}{\mu\phi_s^2}$ can be determined empirically by estimating equation (13). The determination takes advantage of the fact that each factor on the right-hand side of eq. (13) - with the exception of f_{ijs} - varies either with i or j , but never with both simultaneously. This

⁵ In primary and gynaecological care, the reported number of contacts includes those with patients on holiday, working abroad or in similar circumstances (Czihal et al. , 2012). Since the patient's place of residence would be misleading in such cases, the 5% of the contacts with GPs and gynaecologists at the most extreme distances (about 100 km and more) were removed from the data.

allows us to estimate eq. (13) by using dummy variables for the locations i and j (θ -fixed effects) and by running a Poisson regression,

$$\ln(E[x_{ijs}]) = \theta_{is} + \theta_{js} + \theta_s \cdot d_{ij}^2. \quad (29)$$

The calibrated value $\hat{\phi}_s$ is computed from the estimated value $\hat{\theta}_s$ after setting $\frac{1}{\mu} = 1/3$. The resulting $\hat{\phi}_s$ can be interpreted as the upper bound of service providers' catchment areas. Less than 5% of billed cases come from outside this range (Chilès and Delfiner, 1999).

4.4 Measuring access to care

Figures 1 to 3 below graphically contrast the FCA method with the gravity model in their measurement of access to care in Germany. They visualize the access to GPs, to psychotherapists (as an example of ambulatory specialists), and to hospitals, respectively. Finally, Figure 4 is the attempt to visualize an aggregation over all access indices. For the purpose of aggregation, the specialties have been weighted by their share in billed cases.

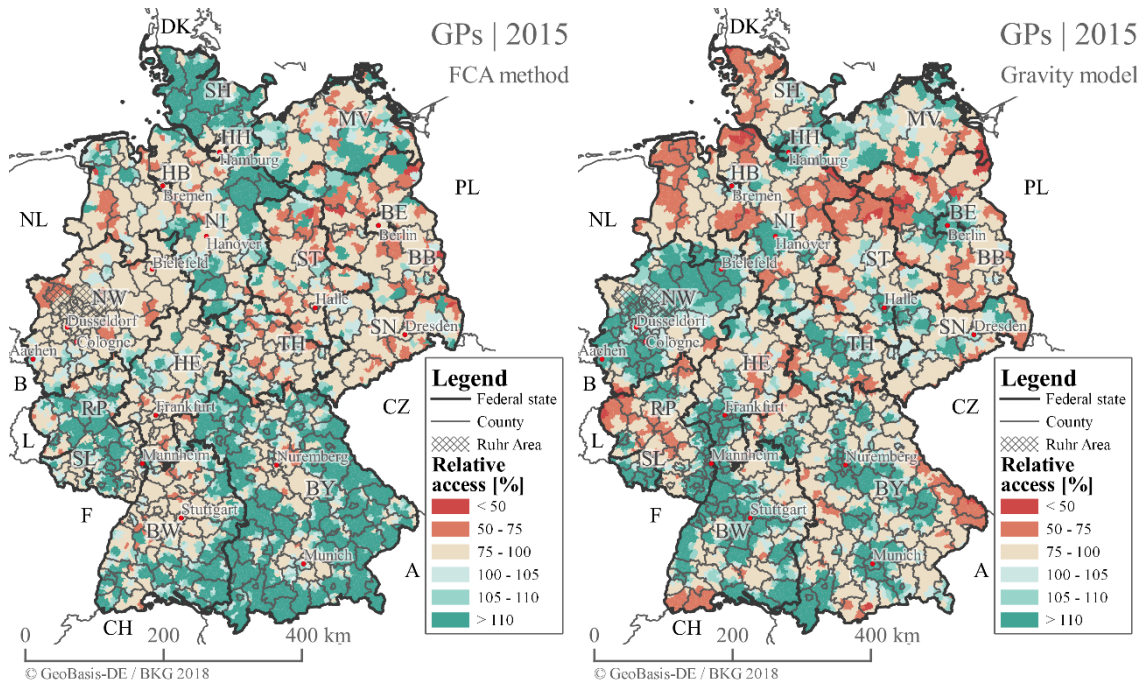


Figure 1: Access to GPs⁶ (left: FCA method; right: gravity model)

⁶ Country codes: A: Austria, B: Belgium, CH: Switzerland, CZ: Czech Republic, DK: Denmark, F: France, L: Luxembourg, NL: Netherlands, PL: Poland; federal state codes: BB: Brandenburg, BE: Berlin, BW: Baden-Württemberg, BY: Bavaria, HB: Bremen, HE: Hesse, HH: Hamburg, MV: Mecklenburg-Western Pomerania,

The first striking impression created by these maps is the great difference in colouring. The regional identification of under- and oversupply seems to be almost contrary. According to the gravity model (right-hand side), larger cities enjoy better access to general practitioners (GPs) than rural communities, just as one would expect. Metropolises such as Hamburg, Munich, Stuttgart and the agglomerations of the Rhine-Main⁷ and Rhine-Ruhr⁸ areas are shown to enjoy particularly high levels of access. The FCA measurement suggests the opposite. The Rhine-Ruhr area and the large cities of Berlin, Munich and Hamburg are identified as regions of more or less severe undersupply. This is implausible and can only be explained by the FCA method's failure to capture spatial direction in the exchange of services (Bauer et al., 2018). The cities and regions referred to above are known for their high capacities and for serving as centres of medical care for surrounding areas (Czihal et al., 2012). The FCA's symmetrical distancing function does not do sufficient justice to the different roles played by cities and their environs. To illustrate this deficiency, let us look at a stylized model of two geographical entities. One of these, denoted by *A*, has all the features of a large city, i.e. a high supply of services and a population which can satisfy its demand locally. The second, denoted by *B*, has a small population, a low supply of services and many people demanding medical treatment in *A*. Because of the spatial symmetry, the FCA method would indicate an equal level of access to care at both places. By contrast, the gravity model would suggest that *A* is better served than *B*, as a higher proportion of local demand is satisfied locally. This theoretical conclusion is confirmed by the real examples of the Rhine-Main and Rhine-Ruhr agglomerations, which do indeed have high inflows and low outflows of demanders for health care (Sundmacher et al., 2018). Another insight conveyed by the gravity model map is that the access to GPs is lower in eastern than in western Germany. This difference mainly results from the spatial distributions of morbidity and deprivation, which both impact strongly on the demand for GPs (Table 2) and are considerably higher in eastern Germany and along the Czech border.

NI: Lower Saxony, NW: North Rhine-Westphalia, RP: Rhineland Palatinate, SH: Schleswig Holstein, SN: Saxony, SL: Saarland, ST: Saxony-Anhalt, TH: Thuringia.

⁷ The metropolitan area around the city of Frankfurt, including Mainz, Wiesbaden, Worms, Darmstadt, Aschaffenburg.

⁸ The Ruhr area plus the cities of Cologne, Düsseldorf and Bonn. In figures 2 and 3 the Ruhr is marked by hatching.

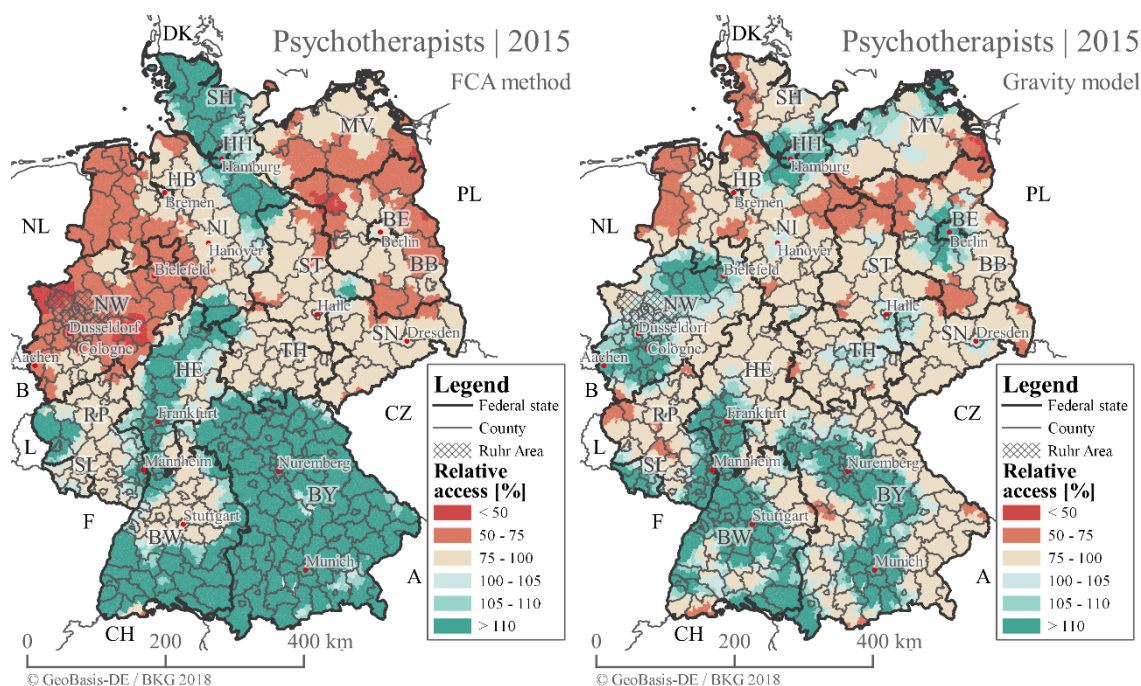


Figure 2: Access to psychotherapists⁵ (left: FCA method, right: gravity model)

Interpreting the differences and similarities between the competing measurements is more difficult in the case of psychotherapy. Explanations based solely on the urban-rural contrast are unconvincing. There are cities such as Hamburg and Munich and the Rhine-Main agglomeration where both measurement methods indicate high levels of access. At the same time, the results for regions like Berlin and the Rhine-Ruhr agglomeration are strongly divergent. In such cases, it is necessary to decide to which measurement to attribute the higher plausibility. An argument which speaks in favour of the gravity model relies on the observation that the demand for psychotherapists increases with income. Since income per capita is below the German average in Berlin and the Ruhr area, it can be assumed that the demand for psychotherapists is relatively low in these areas, allowing access to be high despite only moderate supply. This interpretation is in line with current German capacity planning, which postulates a sufficient level of access to psychotherapists in the Rhine-Ruhr area.

The maps for the other nine ambulatory specialties are not shown here, as the deviations between the competing measurements are smaller. However, the geographical patterns are strikingly similar for all ambulatory specialties, with indices highly correlated at values of 0.75 to 0.96. This means that the access situation at any particular location is either good or bad across all medical specialties.

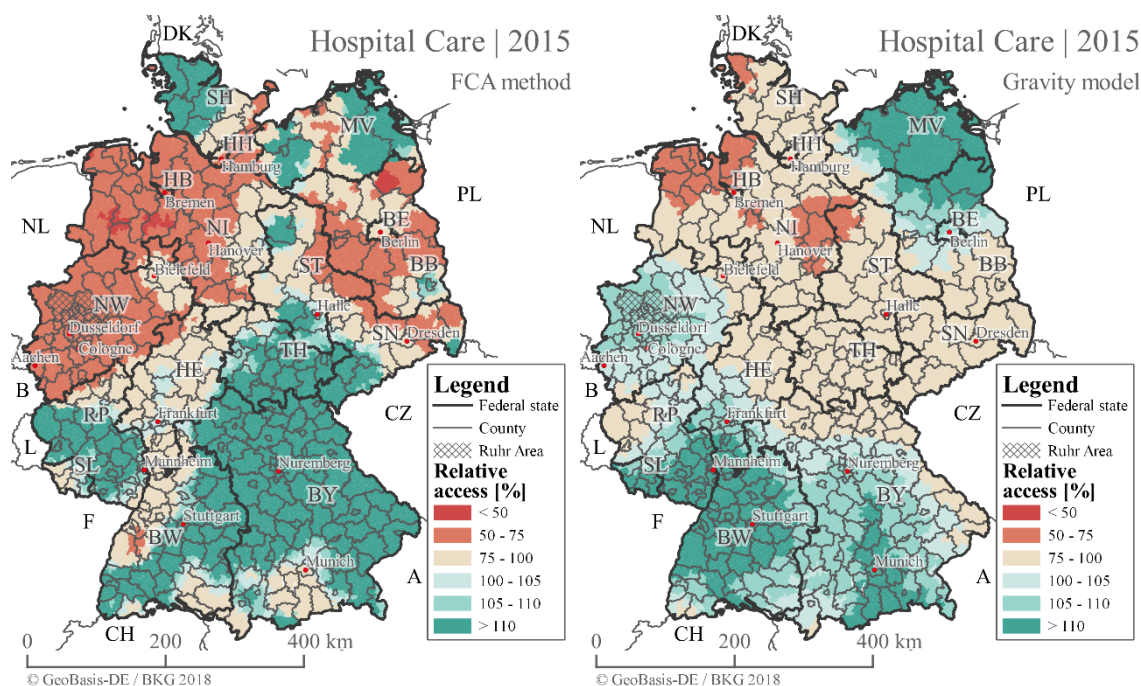


Figure 3: Access to hospitals⁵ (left: FCA method, right: gravity model)

The maps for hospital care (Figure 3) are similar to those for psychotherapy (Figure 2), with similar differences between the measurement methods. On the one hand, there are areas, mainly in the south and the north-east, for which both measurements indicate good access to care. On the other hand, there are areas like the Rhine-Ruhr agglomeration for which the measurements strongly diverge. Here again, more speaks for the gravity model than for the FCA method. The Ruhr area is known to have one of the highest bed densities in Germany, a clear oversupply confirmed by a recent study of German hospital capacities by Loos et al. (2019). The overall picture shown by the gravity-based measurement is that rural areas tend to be undersupplied and urban areas oversupplied, a conclusion also drawn by other studies (Sundmacher et al., 2018). Exceptions to this rule are mainly found in the north. The city of Hamburg seems to suffer from undersupply whereas the opposite holds for parts of Mecklenburg-Western Pomerania and Brandenburg. The oversupply in hospital care in the north-east is best explained by the low population density in this area.

The correlation between the ambulatory and hospital sectors is positive and ranges from 0.34 (hospitals to GPs) to 0.48 (hospitals to urologists). Thus, poor hospital access often goes hand in hand with poor ambulatory access.

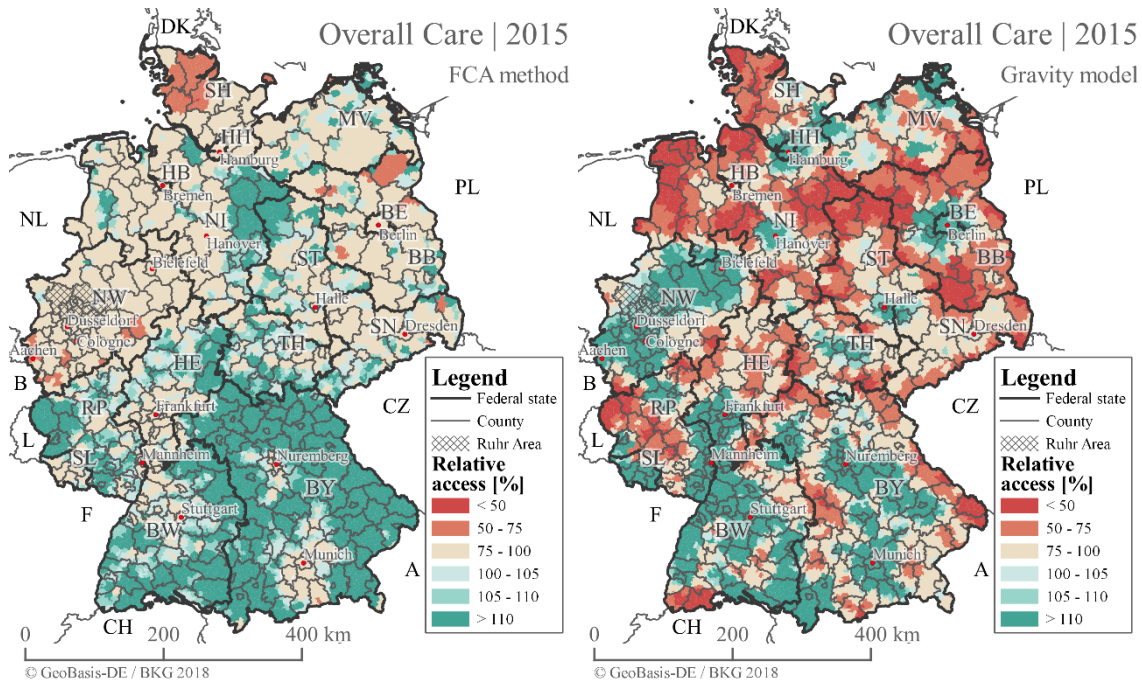


Figure 4: Overall access to care⁵ (left: FCA method, right: gravity model)

The differences in the two competing measurements of access to care become particularly clear when aggregating over the various access indices (Figure 4). Obviously, the measurements produce maps which have little in common. There are few areas for which the two measurements do not indicate opposing situations of care. The general view is in Germany that rural regions tend to be undersupplied and cities oversupplied. The gravity model is more in line with this view. However, the basis of this statement is a qualitative comparison. Therefore, a quantitative comparison is added in the Supplement.

Of course, as has already been pointed out, the original FCA method no longer reflects the state of the art, and using the iFCA method Bauer and co-authors have published maps of Germany that are quite close to those shown here for the gravity model. This applies to GPs (Bauer et al., 2018), gynaecologists (Bauer et al., 2017a), and orthopaedists (Bauer et al., 2017b). However, though the similarity of the maps is remarkable, there remain notable differences in detail. One striking example is the Allgäu Alps in the south of Bavaria. Another is the area around the cities of Greifswald and Stralsund in the far north-east. Here the picture of access to GPs drawn by the gravity model and even by the FCA method is much more positive than the one found in Bauer et al. (2018).

For reasons explained in Section 3.5, no notable differences are to be expected between the gravity-based measurement of access and measurements based on the approach of Bikker and

de Vos (1992). To the best of our knowledge, Bikker and de Vos's method has not been applied to Germany before.

5. Projecting health capacities

5.1 Estimating the demand for health care services

A planning model is by nature forward-looking and needs to be backed up with appropriate forecasts, such as demographic projections (Ono et al., 2013). Past take-up of services is not an undistorted indicator of future demand. For example, account must be taken of such factors as supply-induced demand, barriers to ready access for certain groups of patients and the effects of technical progress, to name but a few. Normative decisions need to be taken as to what part of past take-up can be regarded as relevant for planning and consideration given to how this proportion may change over the course of time (Fleurbaey & Schokkaert, 2009). Similar objections have been raised to the devising of funding allocation models, e.g. risk adjustment models. In practice, these objections are met by using regression models working only with factors relevant for equalization (Ellis et al., 2018). Following this example, we assume that the planning authority has decided that compensation should be made for differences in demand per capita, D_{is}/I_i , that are caused by differences in morbidity and socio-economic status.

We follow the standard procedure and use generalized linear modelling to predict the demand for health care services (Sundmacher et al., 2018). The independent variables suggested by eq. (8) are the access to care, β_{is} , and the population adjustment factor, $\mu_i y/T_i$. The latter is proxied by (a) a morbidity risk score (MRS_i), (b) urbanity (urb_i) measured by the density of population, and (c) a socio-economic (deprivation) risk score (IMD_i). Population size, I_i , serves as the offset variable.

Using β_{is} as an independent variable raises the problem of potentially biased estimation. Equations (13) and (14) make abundantly clear that this variable is not exogenous. Since endogeneity would not be a problem if the FCA method were applied, it makes sense to use the FCA method to first estimate counterparts of β_{is} and then use the predicted values as instruments in a generalized moments (GMM) estimation of the following equation:

$$\begin{aligned} \frac{D_{is}}{I_i} = & \exp(\gamma_{D,s} + \gamma_{MRS,s} \cdot \ln(MRS_i) + \gamma_{ur,s} \cdot \ln(urb_i) + \gamma_{IMD,s} \cdot \ln(IMD_i)) \\ & + \gamma_{\beta,s} \cdot \ln(\beta_{is})) \cdot u_{is} \end{aligned} \quad (30)$$

where u_{is} is the error term with the expected value of one. The coefficients, γ_{\cdot} , are the parameters to be estimated. *A priori* the parameter $\hat{\gamma}_{\beta,s}$ can be expected to be non-negative. A large value would indicate a strong impact of access to care on demand. *Vice versa*, a statistically insignificant value would indicate the lack of such an impact, which can be interpreted as indicating the achievement of equality in access.

5.2 Regression results for catchment areas and medical demand

Table 2 shows the regression results for catchment areas and the demand for medical treatment based on equations (29) and (30), respectively. Due to the large number of observations, the standard errors are so small that they are only shown in the supplementary material.

Table 2: Regression results for catchment areas and the demand for medical treatment

	(32)	Demand for treatment (34)				
	Catchment area	Constant	Morbidity	Deprivation	Urbanity	Access index, $\hat{\beta}_{is}$
	$\hat{\phi}_s$	$\hat{\gamma}_{D,s}$	$\hat{\gamma}_{MRS,s}$	$\hat{\gamma}_{IMD,s}$	$\hat{\gamma}_{urb,s}$	$\hat{\gamma}_{\beta,s}$
General Practitioners	23.64	0.95***	0.31***	-0.01	0.01	-0.01**
Ophthalmologists	36.89	-1.48***	0.62***	-0.12***	0.04***	0.14***
Surgeons	39.91	-3.62***	0.18***	0.52***	0.11***	0.04***
Gynaecologists	30.12	-0.89***	0.02	-0.04***	0.03***	0.11***
ENT-specialists	37.28	-2.21***	0.29***	-0.16***	0.07***	0.21***
Dermatologists	47.45	-1.93***	0.22***	-0.09***	0.06***	0.34***
Paediatricians	35.99	-1.82***	-0.20***	-0.38***	-0.02***	0.25***
Neurologists	43.77	-3.14***	0.55***	-0.03***	0.08***	0.31***
Orthopaedic specialists	38.11	-1.93***	0.68***	-0.50***	0.08***	0.22***
Psychotherapists	56.46	-3.67***	-0.90***	-0.25*	0.11***	0.43***
Urologists	49.09	-2.66***	1.08***	-0.19***	0.07***	0.11***
Hospitals	60.00	-1.74***	0.60***	0.44***	0.03***	0.00

Note: Catchment areas in minutes; population as offset; standard deviations in supplementary material; *** $p(z) < 0.01$; ** $p(z) < 0.025$; * $p(z) < 0.05$; The size of the catchment area of hospitals is set at 60 minutes. An estimation is not possible as data is only available at the district level.

At roughly 24 minutes, the measured catchment area is smallest for GPs, followed by gynaecologists with 30 minutes, paediatricians with 36, and ophthalmologists with 37 minutes. At 56 minutes the measured catchment area is largest for psychotherapists, followed by urologists (49 minutes). As explained in Section 4.3 this implies that a GPs practice at a distance of 20 minutes is only visited in 20% of cases. By contrast, similarly distant paediatricians and psychotherapists are visited in 40% and 70% of cases, respectively. With a

95% bound applied to distance weighting, the measured size of catchment areas largely confirms the values in Fülöp et al. (2011). Significant differences only exist for urologists - 49 here vs. 31 minutes in Fülöp et al., (2011) - and psychotherapists (56 vs. 62 minutes)⁹.

The estimated partial elasticity of demand, $\hat{\gamma}_{MRS,s}$, varies between -0.90 and 1.08. This means that in most specialities increases in the morbidity risk score are estimated to induce sub-proportional increases in the demand for treatment. The demand for urologists is the only one to increase over-proportionally. The effects for paediatricians and psychotherapists are actually negative, while the effect on gynaecologists is insignificant.

The partial elasticity of deprivation, $\hat{\gamma}_{IMD,s}$, varies between -0.50 and 0.52. Individuals living in regions with deprivation one percent above the national average visit surgeons (0.52%) and hospitals (0.44%) more often (*ceteris paribus*), consult GPs equally frequently and seek out all other specialists less frequently. These results confirm the findings of Kopetsch (2007a) and Ozegowski and Sundmacher (2014).

With the exception of GPs and paediatricians, the demand for medical treatment increases with urbanity. The stronger demand of urban residents is in line with Ozegowski and Sundmacher (2014).

The key result in Table 2 is the effect of access to health care on the demand for treatment. With the exception of GPs and hospitals, the access elasticity of demand, $\hat{\gamma}_{\beta,s}$, is significantly positive throughout, ranging between 0.04 and 0.43. An increase of one percent in accessibility, caused for instance by shortened travel time or increased local supply, increases the use of specialist care by an average of 0.19%. By contrast, the utilization of GPs declines by 0.01% and the effect on hospitals is insignificant. Such results are in line with Salm and Wübker (2020) who suggest interpreting them as evidence for the claim that regional variation in the utilization of GPs must be explained by factors of demand such as demographics and other patient characteristics. By contrast, supply factors primarily explain variation in the utilization of specialist care. The positive effect for specialists suggests that better access creates higher demand, just as one would expect¹⁰. The non-positive effects for GPs and hospitals, on the other hand, need some explanation. Here it is important to bear in

⁹ The differences can be explained by the more comprehensive and more recent data in the present study and by counting multiple visits to the same doctor as separately. Fülöp et al. (2011) base their analysis on selected German states and on data from 2010 counting multiple visits as one. As to psychotherapists, the difference could also result from an increase in their number.

¹⁰ For a theoretical discussion see Andersen (1995), and for empirical results see Augurzky et al. (2013); Kopetsch and Schmitz (2014) and Büyükdurmus et al. (2017).

mind that the access indices of GPs and hospitals are positively correlated with those of specialist care. In other words, high specialist coverage tends to go hand in hand with a high level of primary and hospital care capacities. Taking this positive correlation into account, the negative effect which the access to GPs has on the demand for treatment indicates that such consultations often serve as a substitute for specialist treatment. Overall better access to health care triggers stronger demand for ambulatory specialists, which in turn absorbs the demand for GPs (see also Büyükdurmus et al., 2017). That demand for hospital beds remains steady regardless of accessibility is plausible, since German hospital care is highly specialized and usually involves surgery.

5.3 Predicting the future demand for health care

Future demand is theoretically determined by assuming equal access, $\bar{\beta}_s = \bar{\beta}_{is}$, and inserting predictions of future population, morbidity, and socio-economic status into the estimated eq. (30), giving $\bar{D}_{is} = \hat{d}_{is} \cdot \bar{\beta}_s^{\hat{\gamma}_{\beta,s}}$. In this paper, however, the focus is not on a quantification of future demand but on methodological issues, the aim being to illustrate capacity planning based on the gravity model. For this reason, for morbidity, characteristics of population, the working hours of ambulatory physicians,¹¹ and the occupancy rates of hospitals¹² we take the values in the base year 2015. The only difference between the “predicted” and the measured demand then comes from replacing the estimated access indices with their target values, $\bar{\beta}_s$. In Section 3.4 we argued that the choice of $\bar{\beta}_s$ is a political one. In what follows, we simply set $\bar{\beta}_s = 1$ (for all s). The differences between the measured and predicted demand obtained this way are small and not worth reporting in detail. The maximum is 8.6% for psychotherapists and the total difference is -0.6%.

5.4 Moving from predicted demand to projected capacity imbalances

The predicted demand, $\bar{D}_{is} = \hat{d}_{is} \cdot \bar{\beta}_s^{\hat{\gamma}_{\beta,s}} = \hat{d}_{is}$, is set equal to its theoretical specification in eq. (9), thus yielding

$$\vartheta_s^\sigma \cdot \bar{I}_i = \hat{d}_{is} \cdot \bar{\beta}_s^{\hat{\gamma}_{\beta,s}-1} = \hat{d}_{is}. \quad (31)$$

¹¹ The data were collected at a time when doctors were required to allot a minimum of 20 hours per week to consultations for SHI patients. Since then the minimum has been increased to 25 hours (§ 17 Abs. 1a Bundesmantelvertrag-Ärzte).

¹² The occupancy rate of hospital beds is the object of capacity planning at the state level. Over the last ten years, it has remained fairly constant at 77.2%.

By inserting eq. (31) into eq. (11), relying on equations (6) and (12) for planned variables, and assuming \bar{L}_{js} to be exogenously fixed, we obtain the number of doctors needed to meet the predicted demand as a function of the planned access which we have assumed to be one:

$$\bar{A}_{js} \cdot \bar{L}_{js} = \bar{\beta}_s^{\hat{\gamma}_{\beta,s}} \cdot \bar{\tau}_{js}^{1-\sigma}(1) \cdot \sum_i \hat{d}_{is} \cdot f_{ijs}^{-\sigma} = \bar{\tau}_{js}^{1-\sigma}(1) \cdot \sum_i \hat{d}_{is} \cdot f_{ijs}^{-\sigma} \text{ for all } js \quad (32)$$

For details of the derivation see the Appendix. $\bar{\tau}_{js}^{1-\sigma}(1)$ is the solution of the system of equations, $\sum_j \bar{\tau}_{js}^{1-\sigma} \cdot f_{ijs}^{-\sigma} = 1$ for all is .

Table 3 compares required and available treatment capacities. The main result is that there are overcapacities in all specialities. The deficits occurring in rural areas and minor towns are more than compensated for by overhangs in urban areas. The greatest imbalance of 39% is to be found among psychotherapists,¹³ although major spatial imbalances also apply to dermatologists (29%), ENT-specialists (28%) and neurologists (26%). The numbers of surgeons and orthopaedic specialists are those most in need of reduction, whilst GP and hospital bed numbers require the least adjustment.

¹³ The figure of 39% is computed as the sum of absolute deviations divided by the number of available doctors, $(5,178+1,526+2,375)/23,170$.

Table 3: Predicted demand and projected capacity imbalances

Speciality	Capacities in urban areas			Capacities in minor towns			Capacities in rural areas			Overall capacities		
	Required	Available	Δ	Required	Available	Δ	Required	Available	Δ	Required	Available	Δ
General Practitioners	19,722	24,520	-4,798	17,089	15,817	1,272	13,140	11,291	1,849	49,951	51,628	-1,677
Ophthalmologists	1,970	2,695	-725	1,709	1,534	175	1,344	1,050	294	5,023	5,279	-256
Surgeons	1,359	1,872	-513	1,232	1,196	36	992	813	179	3,583	3,880	-297
Gynaecologists	3,702	5,142	-1,440	3,151	2,804	347	2,402	1,803	599	9,254	9,748	-494
ENT-specialists	1,459	2,115	-656	1,224	1,068	156	952	725	227	3,634	3,910	-276
Dermatologists	1,164	1,727	-563	1,048	933	115	805	549	256	3,016	3,209	-193
Paediatricians	2,090	2,845	-755	1,795	1,693	102	1,368	1,075	293	5,254	5,613	-359
Neurologists	1,769	2,576	-807	1,524	1,374	150	1,195	879	316	4,488	4,829	-341
Orthopaedic specialists	1,961	2,812	-851	1,687	1,553	134	1,310	1,007	303	4,957	5,370	-413
Psychotherapists	8,521	13,720	-5,199	7,735	6,203	1,532	5,728	3,342	2,386	21,983	23,264	-1,281
Urologists	999	1,337	-338	881	793	88	709	517	192	2,589	2,644	-55
Hospitals	235,557	307,828	-72,271	224,718	192,688	32,030	175,800	135,891	39,909	636,075	636,407	-332

Note: For the measurement of demand see Table 1. Capacities are measured by the number of doctors in full-time equivalents with the exception of hospitals, which are measured by the number of beds. $\Delta \equiv$ required – available.

The gravity-based analysis in Table 3 suggests that health equity could be dramatically improved by reallocating capacities from urban to other areas. Indeed, it calls current German practice in capacity planning into question by highlighting its apparent failure to identify the true imbalances in health care access. There are three possible explanations for this failure. Firstly, with the exception of psychotherapists (data from 1999), current capacity planning is largely based on outdated doctor-population ratios dating from 1990. Moreover, the target ratios are not derived from the needs of the population, however these are measured, but primarily serve the purpose of limiting the growth in the number of doctors. Secondly, doctors' labour productivity has increased less than the number of medical interventions. As a result, the average doctor bills 6% fewer cases per year than ten years ago (KBV, 2020). The third reason is that the present system of capacity planning aims less at satisfying medical demand than at covering geographic space.¹⁴

No statutory target figures exist for hospital planning but our gravity-based analysis suggests the need for a substantial reallocation of capacities. There are two possible counterarguments to this conclusion. Firstly, hospitals are not easily set up or closed, investments are largely irreversible, and due account must be taken of the proximity of other facilities and institutions such as universities. Secondly, the advantage of easy access needs to be traded off against the economies to be achieved by centralisation. Indeed, many specialized treatments can only be safely performed if they are conducted frequently enough for the staff involved to acquire and maintain sufficient experience.¹⁵ Nevertheless, gravity-based planning should help to identify imbalances in the spatial allocation of hospital capacities.

6. Conclusions

This paper presents a gravity model designed to improve health care capacity planning. The decisive advantage of the approach is methodological. Supply and demand for medical services are modelled by using a general equilibrium framework grounded in choice theory, and indices of access to health care are directly derived from this model. As eq. (15) shows, these indices are inversely proportional to the product of two geometric means. One is taken

¹⁴ Only recently efforts have been made to take greater account of medical needs in planning by setting targets for accessibility. Sundmacher et al. (2018) derive and propose specific targets for ambulatory care and their proposals have now been incorporated into the statutory capacity planning guideline (§35 (2) Fn. 2 Bedarfsplanungs-Richtlinie). These targets are largely confirmed by the present study in terms of both direction and absolute extent.

¹⁵ For this reason Germany's statutory health insurers may refuse to reimburse certain costs of treatment if the billing hospital has not performed a minimum number of such interventions in a given period. See § 136b (2) SGB V.

of cross-border supply ratios and the other of distances. In a comparison, the access index derived from the floating catchment area (FCA) method proves to be a special case (Proposition 2) in the sense that it omits certain adjustments of the population for morbidity and distance suggested by the gravity model. The shortcomings of the FCA method have been a frequent topic in the literature and a number of attempts have been made to correct them by modifying the metric of the population's care potential (e.g. Delamater, 2013). However, these attempts suffer from being largely *ad hoc* and supply-biased.

The present model builds on Bikker and de Vos (1992), whose idea it was to explain supply and demand linkages for inpatient hospital care by a gravity equation. Again, the advantage of our approach lies in its choice-theoretic grounding. As explained in Section 3.5, capacity planning requires full equilibrium modelling to account for interregional and intersectoral supply and demand linkages.

The merits of the present approach are illustrated by applying it to the German health system. Access to care is measured for GPs, hospitals, and ten ambulatory specialities and revealed to be particularly low in wide areas of Germany's north-east across almost all medical specialities.

Finally, in Section 5 we compare the actual capacities for health care with those needed if equality of access is to be achieved. Confirming findings in the literature, we show that there is a clear urban-rural divide. Rural areas and minor towns suffer from undersupply whereas urban areas enjoy considerable oversupply. Aggregating deficits and overhangs, overcapacities are shown to exist in all specialities. We also find that the population of socially deprived areas is more likely to take advantage of hospital care than of treatment provided by ambulatory specialists.

No study of this kind would be complete without caveats. Our greatest reservation concerns the measurement of patient morbidity. The negative coefficients of the *MRS* estimators suggest that they are measuring general chronicity rather than speciality-specific morbidity. This deficiency could be corrected by relying on a more sophisticated morbidity measurement better tailored to medical specialities, as in Sundmacher et al. (2018). A further point of potential criticism is our data, which exclude those who did not consult an SHI doctor in either of the years 2015 or 2016.¹⁶ This may bias the results, especially in regions with a high proportion of patients with private, rather than statutory, health insurance, since their

¹⁶ For further, more minor limitations resulting from the data the reader is referred to Sundmacher et al. (2018).

treatments are not captured by the KBV billing data. A final caveat concerns the size of catchment areas, which may well be underestimated. The reason is that journey times are computed by using the postal codes of billing addresses. The time of journeys within the same postal code is thus zero, a technical simplification which obviously underestimates some actual journey times. Despite all these limitations, however, we strongly believe that our choice-theoretic grounding of the gravity equation and the empirical analysis of the German situation based on this equation make a significant contribution towards bringing health care planning more in line with medical needs.

7. Appendix

Eq. (32) is derived as follows. Equating equations (6) and (12) for planned variables yields

$$\bar{A}_{js} \cdot \bar{L}_{js} = \bar{S}_{js} = \bar{\tau}_{js}^{1-\sigma} \cdot \bar{\alpha}_{js}. \quad (33)$$

Inserting eq. (31) into eq. (11) yields

$$\bar{\alpha}_{js} = \sum_i (\vartheta_s^\sigma \cdot \bar{l}_i) \cdot f_{ijs}^{-\sigma} = \sum_i \hat{d}_{is} \cdot \bar{\beta}_s^{\hat{\gamma}_{\beta,s}-1} \cdot f_{ijs}^{-\sigma} = \bar{\beta}_s^{\hat{\gamma}_{\beta,s}-1} \cdot \sum_i \hat{d}_{is} \cdot f_{ijs}^{-\sigma}. \quad (34)$$

As shown in Section 3.4,

$$\bar{\tau}_{js}^{1-\sigma} = \bar{\tau}_{js}^{1-\sigma}(\bar{\beta}_s) = \bar{\beta}_s \cdot \bar{\tau}_{js}^{1-\sigma}(1). \quad (35)$$

Using equations (34) and (35) and insertion into eq. (33) yield eq. (32). \square

8. Supplement: A quantitative assessment of the gravity-based measurement of access to care

In Section 4.4, we used the gravity model to illustrate access to health care using Germany as an example. It became clear that the gravity model paints a much more convincing picture than the FCA method. One might, however, object that the FCA method does not reflect the state of the art and that the ranking is only based on a qualitative comparison of maps. To meet the second objection, we shortly examine how well the gravity model is able to identify structural supply deficit.

A standard indicator of structural supply deficit is the number of hospital cases with Ambulatory Care Sensitive Conditions (ACSC) (Ansari et al., 2006). Such cases could have been avoided by appropriate ambulatory treatment. Sundmacher and Kopetsch (2015) show

that poor access to ambulatory care has an increasing effect on the number of ACSC cases, just as particularly good access to hospital care is expected to elicit ACSC cases. Against this background, one might wonder which access measurement is better suited to estimate the number of ASCS cases. To investigate the relationship between ASCS cases and access indices, we estimate the following equation for both the gravity model and the FCA method:

$$\frac{E[ASCS_i | IMD_i, SES_i, \beta_{is}]}{I_i} = \exp[\gamma_{ACSC} + \gamma_{ACSC,MRS} \cdot \ln(MRS_i) + \gamma_{ACSC,IMD} \cdot \ln(IMD_i) + \gamma_{ACSC,GP} \cdot \ln(\beta_{i,GP}) + \gamma_{ACSC,specialist} \cdot \ln(\beta_{i,specialist}) + \gamma_{ACSC,Hospital} \cdot \ln(\beta_{i,Hospital})]$$

The variable *MRS* denotes a morbidity risk score chosen to proxy morbidity, μ_i , and *IMD* denotes a deprivation risk score chosen to proxy socio-economic status. *IMD* stands for the index of multiple deprivations proposed by Kroll et al. (2017). It is obtained by aggregating German indices of education, income, and social disadvantage. By contrast, *MRS* is estimated by running an OLS regression in which age-gender (5-years) and disease dummies (according to Elixhauser et al., 1998) are used as independent variables and in which the dependent variable is expressed in units of time and computed by relying on the Uniform Assessment Standard (Einheitlicher Bewertungsmaßstab, EBM, defined by § 87 sect. 2 SGB V). The *MRS* determined in this way is a valid instrument for modelling differences in morbidity (Göpffarth et al., 2016). As to the ASCS definition, we rely on the one suggested by Sundmacher et al. (2015) for Germany. The variables $\beta_{i,s}$ represent the indices of access to (a) GPs, (b) ambulatory specialists when aggregated, and (c) hospitals, respectively.

Table 4: Regression results (ACSC cases)

	Gravity model			FCA		
	γ_i	std. deviation	p-value	γ	std. deviation	p-value
intercept	-4.78	0.16	0.00	-5.44	0.10	0.00
morbidity	-0.00	0.03	0.91	-0.00	0.02	0.87
deprivation	0.71	0.10	0.00	0.92	0.05	0.00
GPs	0.05	0.15	0.75	0.27	0.14	0.02
Specialists	-0.55	0.21	0.01	-0.13	0.07	0.01
Hospitals	0.26	0.10	0.01	0.14	0.06	0.01
R ²	24.23%			22.72%		
AIC	8824.42			8840.32		

The explanatory power of the regression based on gravity indices is the higher one. The Akaike Information Criterion (AIC) is definitely lower and the coefficient of determination is

larger. The difference in R^2 is significant. The estimated coefficients associated with the gravity model are absolutely larger than those associated with the FCA method ones for specialists and hospitals, as can be seen by comparing the regression coefficients in the above Table. The result confirms the expectation that ASCS cases are supply-induced (hospital) and that access to specialists exerts the main (negative) impact. The result is well in line with Sundmacher and Kopetsch (2015).¹⁷ Thus, we conclude that the gravity model seems to be better than the FCA method in identifying structural supply deficits such as high numbers of ASCS cases.

9. References

- Andersen, R., 1995. Revisiting the Behavioral Model and Access to Medical Care: Does it Matter? *Journal of Health and Social Behavior* 36, 1–10.
- Anderson, J., 1979. A Theoretical Foundation for the Gravity Equation. *The American Economic Review* 69, 106–116.
- Anderson, J., 2011. The Gravity Model. *Annual review of Economics*, 133–160.
- Anderson, J., van Wincoop, E., 2003. Gravity with gravitas: A solution to the border puzzle. *The American Economic Review* 93, 170–192.
- Ansari, Z., Laditka, J.N., Laditka, S.B., 2006. Access to health care and hospitalization for ambulatory care sensitive conditions. *Medical Care Research and Review* 63, 719–741.
- Augurzky, B., Kopetsch, T., Schmitz, H., 2013. What accounts for the regional differences in the utilisation of hospitals in Germany? *The European Journal of Health Economics* 14, 615–627.
- Bauer, J., Groneberg, D.A., Maier, W., Manek, R., Louwen, F., Brüggmann, D., 2017a. Accessibility of general and specialized obstetric care providers in Germany and England: an analysis of location and neonatal outcome. *International Journal of Health Geographics* 16, 44.
- Bauer, J., Maier, W., Müller, R., Groneberg, D.A., 2018. Hausärztliche Versorgung in Deutschland – Gleicher Zugang für alle? *Deutsche Medizinische Wochenschrift* 143, e9-e17.
- Bauer, J., Müller, P., Maier, W., Groneberg, D.A., 2017b. Orthopedic workforce planning in Germany - an analysis of orthopedic accessibility. *PLOS ONE* 12, e0171747.
- Bergstrand, J.H., 1985. The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence. *The Review of Economics and Statistics* 67, 474–481.
- Bikker, J.A., de Vos, A.F., 1992. A Regional Supply and Demand Model for Inpatient Hospital Care. *Environment and Planning A* 24, 1097–1116.

¹⁷ The FCA method suggests that any improved access to GPs tends to increase the number of ACSC cases. This result runs against intuition, as it would mean that as the number of GPs increases, patients would more often go to hospital with minor instances. Sundmacher and Kopetsch (2015) do not even include GPs separately in their estimation.

- Büyükdurmus, T., Kopetsch, T., Schmitz, H., Tauchmann, H., 2017. On the interdependence of ambulatory and hospital care in the German health system. *Health Economics Review* 7, 2.
- Chilès, J.P., Delfiner, P., 1999. *Geostatistics: modelling spatial uncertainty*. John Wiley & Sons, Hoboken, New Jersey.
- Czihal, T., von Stillfried, D., Schallock, M., 2012. Regionale Mitversorgungsbeziehungen in der ambulanten Versorgung. *Versorgungsatlas*. Zentralinstitut für die kassenärztliche Versorgung in Deutschland, Berlin.
- Delamater, P., 2013. Spatial accessibility in suboptimally configured health care systems: a modified two-step floating catchment area (M2SFCA) metric. *Health & Place* 24, 30–43.
- Elixhauser, A., Steiner, C., Harris, D.R., Coffey, R.M., 1998. Comorbidity measures for use with administrative data. *Medical Care* 36, 8–27.
- Ellis, R., Martins, B., Rose, S., 2018. Risk adjustment for health plan payment. In: van Kleef, R., McGuire, T.G. (Eds.) *Risk adjustment, risk sharing and premium regulation in health insurance markets. Theory and practice*. Academic Press, London, pp. 1–73.
- Fleurbaey, M., Schokkaert, E., 2009. Unfair inequalities in health and health care. *Journal of Health Economics* 28, 73–90.
- Fülöp, G., Kopetsch, T., Schöpe, P., 2011. Catchment areas of medical practices and the role played by geographical distance in the patient's choice of doctor. *The Annals of Regional Science* 46, 691–706.
- Göpffarth, D., Kopetsch, T., Schmitz, H., 2016. Determinants of Regional Variation in Health Expenditures in Germany. *Health Economics* 25, 801–815.
- Huff, D.L., 1964. Defining and Estimating a Trading Area. *Journal of Marketing* 28.
- KBV, 2020. Behandlungsfälle je Arzt, Berlin. <https://gesundheitsdaten.kbv.de>. Accessed 5 October 2020.
- Kopetsch, T., 2007a. Arztdichte und Inanspruchnahme ärztlicher Leistungen in Deutschland: eine empirische Untersuchung der These von der angebotsinduzierten Nachfrage nach ambulanten Arztleistungen. *Schmollers Jahrbuch: Journal of Applied Social Science Studies* 127, 373–405.
- Kopetsch, T., 2007b. Der Zusammenhang zwischen dem Leistungsgeschehen im ambulanten und stationären Sektor des deutschen Gesundheitswesens. *Jahrbuch für Nationalökonomie und Statistik* 227, 49–64.
- Kopetsch, T., Schmitz, H., 2014. Regional variation in the utilization of ambulatory services in Germany. *Health Economics* 23, 1481–1492.
- Kroll, L.E., Schumann, M., Hoebel, J., Lampert, T., 2017. Regionale Unterschiede in der Gesundheit – Entwicklung eines sozioökonomischen Deprivationsindex für Deutschland. *Journal of Health Monitoring* 2, 103–120.
- Leamer, E.E., Stern, R.M., 1970. *Quantitative International Economics*. Allyn and Bacon, New Brunswick, New York, 209 pp.

- Loos, S., Albrecht, M., Zich, K., 2019. Zukunftsfähige Krankenhausversorgung: Simulation und Analyse einer Neustrukturierung der Krankenhausversorgung am Beispiel einer Versorgungsregion in Nordrhein-Westfalen. Bertelsmann Stiftung, Gütersloh.
- Lowe, J.M., Sen, A., 1996. Gravity model application in health planning: analysis of an urban hospital market. *Regional Science* 36, 437–461.
- Luo, J., 2014. Integrating the Huff Model and Floating Catchment Area Methods to Analyze Spatial Access to Healthcare Services. *Transactions in GIS* 18, 436–448.
- Ono, T., Lafortune, G., Schoenstein, M., 2013. Health Workforce Planning in OECD Countries: A Review of 26 Projection Models from 18 Countries. *OECD Health Working Papers* 62, 1–130.
- Ozegowski, S., Sundmacher, L., 2014. Understanding the gap between need and utilization in outpatient care-The effect of supply-side determinants on regional inequities. *Health Policy* 114, 54–63.
- Paez, A., Higgins, C.D., Vivona, S.F., 2019. Demand and level of service inflation in Floating Catchment Area (FCA) methods. *PLOS ONE* 14, e0218773.
- Salm, M., Wübker, A., 2020. Sources of regional variation in healthcare utilization in Germany. *Journal of Health Economics* 69, 102271.
- Sen, A., 2002. Why health equity? *Health Economics* 11, 659–666.
- Sundmacher, L., Kopetsch, T., 2015. The impact of office-based care on hospitalizations for ambulatory care sensitive conditions. *The European Journal of Health Economics* 16, 365–375.
- Sundmacher, L., Schang, L., Schüttig, W., Flemming, R., Frank-Tewaag, J., Geiger, I., Franke, S., Weinhold, I., Wende, D., Kistemann, T., Höser, C., Kemen, J., Hoffmann, W., van den Berg, N., Kleinke, F., Becker, U., Brechtel, T., 2018. Gutachten zur Weiterentwicklung der Bedarfsplanung i.S.d. §§ 99 ff. SGB V zur Sicherung der vertragsärztlichen Versorgung. order by Gemeinsamen Bundesausschusses (GBA), München, Köln, 810 pp.
- Sundmacher, L., Schüttig, W., Faisst, C., 2015. Ein konsentierter deutscher Katalog ambulant-sensitiver Diagnosen. *Versorgungsatlas* 15.
- Vogt, V., 2016. The contribution of locational factors to regional variations in office-based physicians in Germany. *18726054* 120, 198–204.
- Wan, N., Zou, B., Sternberg, T., 2012. A three-step floating catchment area method for analyzing spatial access to health services. *International Journal of Geographical Information Science* 26, 1073–1089.