

competence in competition + health

**series**

#2016/05

Sarah M. Hofmann and Andrea M. Mühlenweg

**Gatekeeping in German Primary  
Health Care** – Impacts on Coordination  
of Care, Quality Indicators and  
Ambulatory Costs



# Imprint

## EDITOR-IN-CHIEF

Martin Karlsson, Essen

## MANAGING EDITOR

Daniel Avdic, Essen

## EDITORIAL BOARD

Boris Augurzky, Essen

Jeanette Brosig-Koch, Essen

Stefan Felder, Basel

Annika Herr, Düsseldorf

Nadja Kairies-Schwarz, Essen

Hendrik Schmitz, Paderborn

Harald Tauchmann, Erlangen-Nürnberg

Jürgen Wasem, Essen

## CINCH SERIES

CINCH – Health Economics Research Center

Weststadttürme Berliner Platz 6-8

45127 Essen

Phone +49 (0) 201 183 - 6326

Fax +49 (0) 201 183 - 3716

Email: [daniel.avdic@uni-due.de](mailto:daniel.avdic@uni-due.de)

Web: [www.cinch.uni-due.de](http://www.cinch.uni-due.de)

All rights reserved. Essen, Germany, 2016

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.

Sarah M. Hofmann and Andrea M. Mühlenweg

**Gatekeeping in German Primary  
Health Care – Impacts on  
Coordination of Care, Quality  
Indicators and Ambulatory Costs**



Sarah M. Hofmann\* and Andrea M. Mühlenweg†

# Gatekeeping in German Primary Health Care – Impacts on Coordination of Care, Quality Indicators and Ambulatory Costs

## Abstract

*Evaluation studies on gatekeeping in primary care exist for a variety of countries but provide mixed evidence on utilization and quality of care as well as costs. Our study evaluates the German gatekeeping program, based on claims data of a major statutory health insurance company. The panel structure of the data allows controlling for patients' characteristics in the year before opting (or not opting) for a GP contract. In contrast to previous studies we are able to draw on multiple identification strategies. We exploit variation in the regional provision of gatekeeping in an instrumental variable (IV) framework. We also analyze GP fixed effects based on the observation of patients opting for one of two different contracts within the same GP office. We find that the gatekeeping contract yields a somewhat higher coordination of care, improved quality (regarding prevention and avoidance of hospitalization) but also higher ambulatory costs. The effects are largely robust between our identification strategies.*

*JEL Classifications: I10, I11, I13*

*Keywords: primary health care, gatekeeping, health care quality.*

---

**Acknowledgements:** *This independent research was initiated from our previous project on behalf of the HÄVG (German GP association). We are also grateful to the IKK classic (the cooperating health insurance company) for data access and for excellent data support, particularly in coding the cost measures. Furthermore, we are grateful to Martin Karlsson and seminar participants and colleagues at the WifOR institute, the CINCH (University of Duisburg-Essen) and the Institute of Labor Economics at the Leibniz University of Hannover, participants of the German Health Association Conferences (DGGÖ) in Bielefeld and Berlin as well as to participants of the 11th European Health Economics Association Conference in Hamburg for valuable comments on earlier versions of this work. All remaining errors are our own. We are not aware of any conflict of interest.*

*\* WifOR Darmstadt and Universität Duisburg-Essen. **Corresponding Author:** WifOR Darmstadt Rheinstraße 22 D-64283 Darmstadt Phone: +49 (0) 6151 50 155 22 Fax: +49 (0) 6151 50 155 29 [andrea.muehlenweg@wifor.de](mailto:andrea.muehlenweg@wifor.de)*

*† WifOR Darmstadt and Leibniz Universität Hannover.*

## 1 Introduction

In many countries a strong primary care system is considered key to an efficient provision of healthcare (EXPH, 2014). Efficiency considerations refer to the role of general practitioners (GP) who might act as gatekeepers with respect to secondary and specialized care (Kringos et al., 2013; Scott, 2000). Gatekeeping is broadly defined as a system where GPs act as primary contact persons, while patients need their “gatekeeper’s” referral in order to access most specialist care. Although policy makers’ attitudes towards gatekeeping are mixed, these systems are on the rise in many countries (e.g. Abbasi, 2014). In an international overview of health care systems in 18 developed and middle-income countries, Mossialos et al. (2016) report that most of these countries have gatekeeping regulations at least in some regions or under specific care plans. According to this overview, only three out of the 18 countries do not offer gatekeeping programs (i.e. India, Singapore and Sweden in 2015). In the US, primary care doctors act as gatekeepers within some managed care plans (Mossialos et al., 2016). In Europe, gatekeeping systems exist both in tax-funded systems like the UK and Spain and in social insurance systems like in Switzerland, in the Netherlands and in Germany (Garrido et al., 2011). In Germany, gatekeeping contracts are currently broadly introduced and co-exist with standard non-gatekeeping contracts (Brekke et al., 2007, Garrido et al., 2011).

To our knowledge, to date there is limited evidence on the quality and cost impacts of gatekeeping. Existing evaluations of gatekeeping programs are hampered by the fact that patients are solely observed in either the gatekeeping or a non-gatekeeping system. Even if both schemes co-exist within a country, the selection of patients into these schemes needs to be taken into account when analyzing the gatekeeping effects.

In the present paper, we analyze the impact of the German gatekeeping program on coordination of care as well as on quality indicators and ambulatory costs. Our paper adds to the existing literature by drawing on objective quality of service and cost measures from an extensive panel data set. Based on health insurance claims data of a major public health insurance company, we observe all insured individuals in the Southern German state of Baden-Württemberg in the year before and up to three years after the introduction of gatekeeping contracts in 2010. The panel structure of the data allows us to control for patients’ characteristics in the year before opting (or not opting) for a GP contract.

One may doubt whether the previously applied control group approaches manage to control for all factors causing selection to the programs and also being related to the outcome variables of interest. Therefore, in contrast to existing studies summarized below, we draw on multiple identification strategies. We do not solely rely on matching or regression analysis controlling for potential sources of selection bias but also exploit variation in the regional

provision of gatekeeping in an instrumental variable (IV) framework. In this respect, we assume that the regional provision affects individual participation but is not directly related to individuals' health outcomes. In a further empirical strategy, we analyze GP fixed effects based on the observation of patients opting for one of two different contracts within the same GP office. Based on this, we are also able to discuss potential spill-over effects of gatekeeping (on non-participants) within the physician's office.

From a theoretical point of view, GPs' gatekeeping may be counteracting information asymmetries, particularly if GP-ensured continuity of care goes along with improved assessment of patients' health histories and needs (Brekke et al., 2007; Scott, 2000). Gatekeeping is expected to avoid overtreatment and to reduce health care costs.<sup>1</sup> However, inefficiencies may also be enforced if the gatekeeping physicians' objectives diverge from patients' interests (e.g. McGuire, 2000; Scott, 2000): GPs may overtreat due to supplier-induced demand related to profit maximization. On the other hand, undertreatment may occur if contracted GPs are tempted to reduce effort in the absence of potential competitors (Godager et al., 2014; Vedsted and Olesen, 2011; Mariñoso and Jelovac, 2003; Gaynor and Town, 2011 for a general summary of competition effects).

Existing empirical research provides mixed evidence on the impacts of strong primary care in general and of the particular impacts of gatekeeping. Garrido et al. (2011) is a systematic review of the international gatekeeping literature: The authors identify 26 studies with documented effects of gatekeeping on health indicators, measures of health care utilization as well as costs. Most of these studies relate to gatekeeping in the US (i.e. 16 out of 26 studies). The remaining studies are based on data for Switzerland, Denmark, Germany, Scotland and the Netherlands respectively. The majority of studies suggest that gatekeeping is related to lower utilization of health services as well as lower costs. As indicated by Garrido et al. (2011), one drawback of the existing evidence is that "health- and patient-related outcomes have been studied only exceptionally and are inconclusive" (ibid., p. 36). The provided evidence varies in magnitude and also in the direction of effects, especially when patient-related outcomes are considered. Garrido et al. (2011) also report that the existing evidence is based on "limited quality" of research designs (ibid., page 36). Specifically, most studies suffer from multiple sources of bias and fail to take into account patient characteristics such as clinical or socio-demographic information. Only six studies consider morbidity or self-reported health status.

---

<sup>1</sup> Correlative empirical evidence suggests that information on the patients' treatment history and medical conditions may exert positive as well as negative impacts, e.g. due to shorter consultations or "wait and see" behavior (Vedsted and Olesen, 2011; Weiss and Blustein, 1996; Hjortdahl and Borchgrevink, 1991).

More recent studies relate to cross-country-analyses. Comparing primary care systems in European countries, correlative evidence suggests that strong primary care is associated with lower rates of unnecessary hospitalizations and to improved population health (Kringos et al., 2013). Hansen et al. (2015) argue that patients suffering from chronic conditions are more likely to be in good self-rated health in countries with a strong primary care structure. In contrast, Vedsted and Olesen (2011) highlight potential adverse effects of gatekeeping on health care quality and health outcomes. Comparing cancer survival rates in European countries, they find that the first-year relative cancer survival rate is significantly lower in gatekeeping systems. According to Kringos et al. (2013), health expenditures seem to be higher in countries with strong primary care systems. However, Docteur et al. (2003) as well as Gerdtham (1998) document lower expenditures related to strong primary care systems.

A number of recent studies evaluate aspects of the recently introduced German gatekeeping contracts and address the statistical challenge of initial selection of patients into programs. In Germany, the broad introduction of gatekeeping contracts is one result of the *Act on Modernization of the Statutory Health Insurance* which was implemented in 2004 and aims at improving coordination of service provision between health care sectors. Since 2009, all statutory health insurance companies are obligated to offer gatekeeping contracts as an alternative to the standard contracts.

Most of the previous evaluation studies for Germany apply matching techniques to generate control groups based on observables (AQUA, 2013; Ose et al., 2008; Riens et al., 2011; Steiner et al., 2013). Likewise, further studies use regression analysis (Böcken, 2008; Gerlach and Szecsenyi, 2013a/b, 2014, 2016<sup>2</sup>; Schnitzer et al., 2011) controlling for observable characteristics.<sup>3</sup> In summary, findings from these studies are mixed, particularly when considering coordination, medication, cost indicators as well as secondary care indicators. More consistently, some studies suggest that gatekeeping is related to a higher number of GP contacts (i.e. the studies by Gerlach and Szecsenyi 2013a/b, 2014; Steiner et al., 2013). Concerning objective measures of care quality based on claims data, Gerlach and Szecsenyi (2014/16) suggest that gatekeeping patients are more often treated according to observed treatment guidelines for specific chronic conditions. Similarly, patients under gatekeeping more often receive preventive treatment (AQUA, 2013; Gerlach and Szecsenyi, 2014/16) and are less often reported to be subject to ambulatory care sensitive

---

<sup>2</sup> Gerlach and Szecsenyi (2016) does not clearly state the applied method but indicates that results are adjusted for control variables.

<sup>3</sup> The most comprehensive evidence is provided in Gerlach and Szecsenyi, 2013a/b, 2014, 2016. These findings relate to several data waves of a major health insurance company (the "AOK"). Included control variables in these studies are similar to the set of controls used in our study. Also, the data stems from the same Southern German state as the data we use in our study.



hospitalizations as well as severe complications observed within the group of patients suffering from diabetes (Gerlach and Szecsenyi, 2014/16).<sup>4</sup>

Our paper consistently finds that the gatekeeping contract yields a higher coordination of care, improved quality (regarding prevention and avoidance of hospitalization) but higher ambulatory costs. Our findings are in line with evidence provided in the studies by Gerlach and Szecsenyi (2013a/b, 2014, 2016), which provide the most comprehensive analysis of gatekeeping in Germany to date (both concerning outcome measures and control variables).

Our paper proceeds as follows: Section 2 describes the main features of the German gatekeeping contracts. Section 3 introduces our data source and presents descriptive evidence on patients' self-selection into the gatekeeping program. Section 4 presents our identification strategies and provides further detail on the spatial distribution of GP contracts. Section 5 presents the results, which are then further discussed and concluded in Section 6.

## **2 Features of the German gatekeeping contracts**

Health insurance companies in Germany operate within a largely universal public health care system: Individuals are obligated to contract with one of the statutory or private health insurance companies. The statutory health insurance companies' contracts offer identical mandatory services and only differ slightly with respect to membership rates and special services.<sup>5</sup> Since 2009, all statutory health insurance companies are obligated to offer gatekeeping contracts as an alternative to standard health insurance plans. Participation in these gatekeeping programs is voluntary for both practitioners and the insured. Patients may opt for either a gatekeeping or a standard contract. GPs participating in the gatekeeping program are operating under two coexisting schemes: Some of their patients will, while others probably won't, opt for the gatekeeping contracts. This coexistence allows us to directly compare the gatekeeping care scheme to the standard scheme within the same national health care system – and even within GP offices.

In this section, we detail three main features of the German gatekeeping contracts which are relevant to our empirical designs. We address (a) the key aspects related to the coordination of care, (b) GPs' remuneration as well as (c) aspects of the structure dimension of care quality.

---

<sup>4</sup> Further findings from the German gatekeeping literature relate to patients' satisfaction with services and measures of self-assessed health (e.g. Böcken, 2008 and Schnitzer et al., 2011). Results are inconclusive and potentially suffer from reporting biases (also see the discussion in Hofmann and Mühlenweg, 2016).

<sup>5</sup> Self-employed persons and employees who earn above a certain threshold (about EUR 50,000 monthly gross earnings) are free to choose between statutory or private health insurance contracts. In 2011, about 13 % of the working population was insured under private health insurance contracts (Destatis, 2013).

Concerning the coordination of care (a), patients are generally free to consult different GPs as well as specialists under the standard (non-gatekeeping) contract in Germany. Under the gatekeeping contract, GPs act as gatekeepers to specialist care, but the respective patients still have direct access to specific specialists (i.e. to gynecologists, optometrists, dentists, pediatricians as well as in emergency cases).<sup>6</sup> This is similar to the gatekeeping systems in Denmark, Estonia and Poland (EXPH, 2014). In contrast, for example Croatia, the Netherlands, Spain, Slovenia, the United Kingdom apply stricter referral regulations (EXPH, 2014).

A further key feature of the German gatekeeping scheme is that remuneration for participating GPs differs from the standard remuneration scheme (b). For patients under the standard contract, GPs' remuneration is based on a mix of a capitation and a fee-for-service scheme: Services are remunerated separately, while practitioners receive a minor basic rate per patient treated within the quarter. In contrast, remuneration for GP services provided for program participants is capitation based; only few services are being paid for separately.

From a theoretical point of view capitation systems may generally incentivize GPs to attract as many patients as possible while minimizing effort (Scott, 2000; Blomqvist and Leger, 2005; Allard et al., 2011). This might counteract the desired quality effects of gatekeeping. Similarly, one may expect that systems with capitation schemes yield lower per patient expenditures compared to fee-for-service schemes (cf. the cross-country evidence in Gerdtham and Jönsson, 2000). On the other hand, GPs in the German gatekeeping-capitation scheme receive a rather high fixed rate per patient.<sup>7</sup> The relatively generous baseline remuneration structure is supposed to be an incentive for GPs to participate in the program (Hausärzterverband, 2015). Therefore, there might be limited scope for direct primary care cost savings of the gatekeeping/capitation scheme.

Fee-for-services for patients under gatekeeping relate to prevention measures such as check-ups and influenza vaccinations, ambulatory operations and treatments of psychosomatic conditions. Thus, for these specific services, GPs income is *always* directly linked to service provision. This might generally incentivize GPs to provide these treatments. In our empirical analysis we consider outcome variables both related and unrelated to the fee-for-service incentive: This gives some guidance towards the gatekeeping effect interacting with financial incentives (cf. Section 3 and Discussion in Section 6).

---

<sup>6</sup> German gatekeeping contracts imply that patients may see all other medical specialists only based on referrals. However, patients may ignore this regulation and still receive treatment. To date, there are no penalties for not complying with the contract.

<sup>7</sup> The rate amounts to EUR 60 to EUR 300, depending on classification of chronic conditions (Hausärzterverband, 2016).

A further feature of the gatekeeping contracts relates to service standards for participating practitioners (c). Participating GPs are requested to regularly attend advanced training courses on medication and treatment guidelines, to implement up-to-date IT infrastructure, to offer special services such as consultation hours on weekends, and to comply with cost saving requirements when prescribing drugs. These prerequisites probably yield a positive selection of practitioners into the GP program which needs to be taken into account in our empirical identification strategies.

Similar to GPs, patients also self-select into gatekeeping. Better coordination and a potential rise in service quality provide incentives to participate (Hausärzteverband, 2015). The self-selection of patients will be examined further in Section 3. Participation rates of both practitioners and patients in the GP program vary from state to state and among health insurance companies due to the differing timing of the programs' introduction. Overall, by the end of 2014 around 3.5 million members of all statutory health insurance companies (about 5%) participated in the GP program.<sup>8</sup>

### 3 Data

Our analysis is based on claims data from the billing process of the IKK classic, which is a major statutory health insurance company in Germany.<sup>9</sup> The available data set contains all insured residents of the federal state of Baden-Württemberg contacting a GP in the years 2009 to 2013 and with consistent documentation of their health insurance status.<sup>10</sup> The considered health insurance company introduced gatekeeping contracts in October 2010. Thus, the panel data allows observing individuals before and after the introduction of the program. Available individual-level information is related to diagnoses, treatments, operations and medical prescriptions (out-patient as well as in-patient) and ambulatory costs. Furthermore, we observe individual characteristics such as age, sex, zip codes and a proxy for labor force participation.

We restrict our sample to the relevant age group of persons aged 18 and older at the time of their decision to participate in the program. Additionally, we limit the sample of program participants to patients opting for the contract until the third quarter of 2011. Persons opting for a gatekeeping contract later on are excluded from the analysis. This implies that all

---

<sup>8</sup> Information provided by the German GP association (HÄVG).

<sup>9</sup> According to recent numbers, the IKK classic is the sixth largest statutory health insurance company in Germany covering about 3.5 million members (Krankenkasseninfo, 2016).

<sup>10</sup> Our samples only include persons visiting a GP which is crucial in order to control for GP characteristics. In a previous version of our study, we have estimated effects based on a larger sample of patients irrespective of actual GP visits and without controlling for GP characteristics. This specification yields robust results in a propensity score matching framework (results from our project report, available upon request). More generally, the propensity score matching results are robust to the regression results presented in this paper.

patients under gatekeeping have been consistently participating for at least 18 months by the time of our outcome measurement in 2013.

### ***Outcome Measures***

The database allows us to analyze key features and objectives of the gatekeeping contract. Feasible outcome measures relate to (a) the utilization of health care services, (b) the quality of care and (c) ambulatory health care costs. For our analysis, we were granted limited access to the claims data, under strict data security restrictions. Because of the limited time for onsite analyses, we had to decide on a manageable quantity of indicators which are detailed in the following paragraphs.

First of all, with respect to the *utilization of services (a)* and in order to measure whether the gatekeeping program is doing what it is supposed to do, we consider the number of GPs visited within a year. Theoretically, under gatekeeping the number of GPs should always be one. Due to vacation replacement regulations the average gatekeeping patients' number of GPs may be slightly higher. Also, to date there are no penalties for patients not complying with the terms of their gatekeeping contracts. Therefore, it is possible that patients deviate from the gatekeeping premise.

As a further measure of the utilization of health care services (a) we observe the utilization of specialist treatment with and without reported referral as direct indicators for the coordination of care. Specialist visits are observed as treatment cases by medical specialist within one quarter of a year, whereas a treatment case may imply multiple visits to the same specialist:<sup>11</sup> Since 2010, the claims data do not document separate information on each single visit. Also, referrals tend to be underreported because reporting is upon responsibility of the specialist. However, to our knowledge there is no reason for a systematic difference in the documentation practice for program participants and non-participants. Therefore, the numbers of specialist visits with and without referrals are considered to be valid outcome measures (also see Hofmann and Mühlenweg, 2016).

Additionally, we measure coordination of services based on the observation of patients with chronic conditions participating in disease management programs (DMP). Within the statutory health care system, DMPs are offered to patients suffering from diabetes<sup>12</sup>, breast cancer, coronary heart disease (CHD), asthma or chronic obstructive pulmonary disease (COPD). Diagnoses and treatment of all of these conditions except for breast cancer are within the scope of responsibility of the GP. Therefore, we do not consider the breast cancer

---

<sup>11</sup> For treatment without referral also see footnote 6.

<sup>12</sup> Concerning diabetes type I and II, we only observe new registration into the diabetes type II program in our data.

DMP. For the other DMPs, we measure registrations of patients diagnosed with the respective chronic condition after potential enrolment in the gatekeeping contract.

With respect to *health care quality (b)*, we observe three outcome measures. Two of these measures are related to prevention<sup>13</sup>: Particularly, we observe influenza vaccinations in patients aged 65 and older as well as general preventive health “check-ups” in patients aged 35 and older. Concerning influenza vaccinations in patients aged 65 and older, they are considered to be an important preventive measure (see e.g. the *Quality Indicators in Ambulatory Health Care*, AQUA, 2009). Influenza vaccinations have been shown to be associated with a significant reduction in mortality and in the risk of respiratory diseases in the elderly (Gross et al., 1995; Nichol et al., 1998; Wilde et al., 1999). As to the second measure, the general health “check-up” offered to all insured persons aged 35 and older, this aims at detecting chronic diseases which may be asymptomatic, such as diabetes mellitus, cardiovascular diseases and renal diseases. Early detection of these chronic diseases is considered to be crucial for successful treatment (Hoebel et al., 2013). We consider both prevention measures as depicting aspects of the process dimension of care quality (following the quality classification according to Donabedian, 1988)

Besides these measures related to the process dimension of quality, we draw on *Ambulatory Care Sensitive Conditions (ACSH)* as an indicator for the *outcome* dimension of care quality. ACSH are defined as hospitalizations that may be prevented through adequate primary care (see e.g. Harrison et al., 2014). In our analysis, we rely on 19 sub-indications that are considered to be the most relevant ACSH according to the National Health Service (NHS) in England (see Sundmacher and Kopetsch, 2015).<sup>14</sup>

Concerning the cost-effectiveness dimension, we observe ambulatory health care costs. This measure comprises all *billed costs* for primary and secondary care services provided by health professionals outside hospitals and generally covered by the statutory health insurance. The available measure is not necessarily identical to the finally reimbursed costs which may be subject to ex post adjustments of fees provided per service to the physicians in the fee-for-service scheme.<sup>15</sup> Also, costs for medication are not included in this cost measure.<sup>16</sup>

---

<sup>13</sup> We use these measures because to our knowledge these are the only GP-based prevention measures with consistent information for all patients in the claims data.

<sup>14</sup> Sundmacher et al. (2015) develop an ACSH list for Germany. This list had not been published by the time we generated our indicator. Therefore, we rely on the NHS elaboration of relevant indications according to Sundmacher and Kopetsch (2015).

<sup>15</sup> These cost caps are unobserved in our data. This implies that a cost-increasing effect of the gatekeeping program may be underestimated, i.e. the absolute difference in *reimbursed costs* would be higher than the difference *in billed costs*.

<sup>16</sup> A comprehensive analysis of cost effectiveness would require considering cost indicators together with (monetarized) benefits of the gatekeeping program. An extensive cost-benefit-analysis reaches

We expect our quality measures to interact with the financial incentives of the German gatekeeping program in different ways (cf. section 2). Particularly, for influenza vaccinations fee-for-service remunerations for gatekeeping GPs depend on reaching a vaccination threshold of 55 % among their elderly gatekeeping patients. Income-maximizing gatekeepers have a clear incentive to conduct vaccinations among the gatekeeping patients. Therefore, observing higher vaccination rates among program participants might be rather due to financial incentives than to the gatekeeping effect. In contrast, general health check-ups are equally remunerated under the fee-for-service scheme for program participants and non-participants. In this case, one may expect that there is no, or to a lower extent, interaction of the gatekeeping effect with a financial incentive.

There are no specific financial incentives (fee-for-services) related to *Ambulatory Care Sensitive Conditions*. If GPs maximize profits, we would expect that they limit treatment quality under the capitation scheme. In this case, the capitation element of the gatekeeping contracts might counteract the potentially improved quality through better coordination.

Generally, it would be interesting to observe further indicators, especially measures of the *outcome* dimension of quality. In our own previous work, we have discussed a number of potential indicators and their feasibility in claims data (Hofmann and Mühlenweg, 2016). Also, the most recent German evaluation study of Gerlach and Szecsenyi (2016) uses a variety of quality indicators related to chronic conditions. However, quality indicators based on chronic conditions might be considered to be of limited validity: GPs in the gatekeeping program receive higher lump-sum payments for patients they classify as suffering from a specific condition. Because of this clear monetary incentive, we expect GPs diagnosing behavior to differ for patients participating in gatekeeping and for non-participants. As a stylized fact, in our data we observe that less than 10 % of patients in the treatment and control groups (5.6 % and 4.1 % respectively) are classified as suffering from chronic depressions when observed before introduction of the gatekeeping program in 2009. In contrast, in 2013 about 35.0 % of the gatekeeping patients are classified as depressed (while the rate remains more stable in the control group, 6.5 % in 2013). These facts hint to chronic conditions being more likely to be diagnosed in patients suffering from less severe chronic symptoms among participants as compared to the non-participants. In this case, comparisons of clinical endpoints and further complications of participants and non-participants are not considered adequate means to evaluate the gatekeeping program.

### ***Observable characteristics of participants and non-participants***

---

beyond the scope of the present paper, where we only observe a limited number of indicators on costs and quality benefits. Our work can be seen as a first step towards a more comprehensive framework.

Table 1 presents mean characteristics of future participants and non-participants in the gatekeeping program. The respective characteristics are observed prior to the gatekeeping choice (in 2009). All of the presented variables are used as control variables in our econometric analysis. Some of the respective variables also constitute outcome measures when observed in 2013.<sup>17</sup>

Prior to their actual decision to participate in the gatekeeping program, the average future participant already differs substantially from the average non-participant in terms of age, morbidity as well as utilization of health care services (Table 1). This is in line with existing evidence on self-selection of patients into the German gatekeeping program (Freund et al., 2010; Kürschner et al., 2011). Particularly, Table 1 reveals that future program participants are on average about five years older than non-participants; participants are significantly more often diagnosed with severe illnesses such as heart disease, COPD or diabetes. Accordingly, their number of GP and specialist visits is higher (8.3 vs. 6.3 and 4.3 vs. 3.5), and annual ambulatory health care costs exceed those of future non-participants by about EUR 120. Also, future participants participate more often in disease management programs (DMPs). In general, future gatekeeping GPs offer these programs more frequently per se. For future gatekeeping patients, about 86 % of the respective GPs also offer DMPs. For future non-participants, only 67 % of GPs participate in DMPs (see bottom row of Table 1). Future gatekeeping participants' GPs also differ in that they tend to operate in larger offices (mean of 235 vs. 205 patients).<sup>18</sup>

**Table 1: Observable characteristics of future participants and non-participants (2009)**

	Standard care program		Gatekeeping program	
	Mean	(SD)	Mean	(SD)
<i>Socio-demographic characteristics</i>				
Age	47.11	(17.67)	52.37	(16.92)
Female indicator	0.48	(0.50)	0.48	(0.50)
Labor force indicator <sup>A</sup>	0.54	(0.50)	0.48	(0.50)
<i>Morbidity</i>				
Diagnosed with...				
Myocardial infarction	0.01	(0.09)	0.01	(0.11)
Heart disease	0.02	(0.14)	0.04	(0.19)
Peripheral vascular disease	0.02	(0.14)	0.03	(0.17)
Cerebrovascular disease	0.03	(0.16)	0.04	(0.21)
Dementia	0.00	(0.05)	0.00	(0.06)
COPD	0.08	(0.28)	0.12	(0.32)
Connective tissue disease	0.01	(0.12)	0.02	(0.14)
Ulcer	0.01	(0.07)	0.01	(0.09)
Liver disease	0.00	(0.05)	0.00	(0.06)

<sup>17</sup> Thus, including the 2009 values as control variables follows the logic of a difference-in-difference identification strategy.

<sup>18</sup> The number of patients within GP office refers to patients insured with the health insurance company considered in this study and being treated in the respective GP's office in 2009.

Diabetes mellitus type I or II	0.08	(0.28)	0.14	(0.35)
Diabetes mellitus complications	0.01	(0.12)	0.03	(0.16)
Paraplegia	0.01	(0.08)	0.01	(0.08)
Renal disease	0.02	(0.13)	0.02	(0.15)
Cancer	0.03	(0.18)	0.05	(0.22)
Cancer (metastasizing)	0.00	(0.06)	0.01	(0.07)
Liver disease (severe)	0.00	(0.02)	0.00	(0.01)
HIV	0.00	(0.03)	0.00	(0.02)
In nursing care	0.00	(0.07)	0.00	(0.04)
<i>Health care utilization</i>				
Ambulatory health care costs	429.61	(528.34)	550.45	(564.30)
Medication costs	323.55	(1,295.36)	440.67	(1,427.43)
Number of GP visits	6.29	(6.01)	8.30	(6.31)
Number of different GPs	1.49	(0.97)	1.66	(0.98)
Number of specialist visits	3.54	(3.78)	4.27	(4.13)
Number of hospitalizations	0.16	(0.51)	0.20	(0.56)
DMP asthma participation	0.00	(0.06)	0.01	(0.12)
DMP COPD participation	0.00	(0.06)	0.01	(0.12)
DMP diabetes type II participation	0.04	(0.20)	0.10	(0.30)
DMP diabetes type I participation	0.00	(0.03)	0.00	(0.03)
DMP CHD participation	0.01	(0.11)	0.04	(0.20)
<i>GP characteristics<sup>B</sup></i>				
Group practice	0.39	(0.49)	0.41	(0.49)
Number of patients treated	204.55	(136.08)	235.03	(142.66)
Offer of DMP	0.67	(0.47)	0.86	(0.34)
<i>Number of observations</i>	<i>245,246</i>		<i>36,209</i>	

**Note:** <sup>A</sup>The *labor force indicator* refers to being an employee, as only this information is documented in the claims data. <sup>B</sup>GP characteristics refer to the gatekeeping GP for future gatekeeping participants. For non-participants, GP characteristics refer to the patient's most frequently contacted GP (> 50% of contacts).

**Source:** Own calculations based on data for a major health insurance company in the German state of Baden-Württemberg (control variables relate to the year before introduction of the gatekeeping program in 2009).

## 4 Empirical strategies

As illustrated in the previous section, participation in the gatekeeping program depends on individual characteristics that may also influence patients' health outcomes and health care costs. For example, one may expect that older or less healthy patients (i.e. those under gatekeeping) cause higher treatment costs, irrespective of the quality of the potential gatekeeper's services. Simply comparing health outcomes or costs of program participants and non-participants would lead to misleading results. In the following section, we draw on multiple identification strategies addressing different sources of biases in different ways. Particularly, the panel structure of our data that observes individuals over time (and also within GP offices) allows us to apply designs based on a difference-in-difference logic (1), an instrumental variable approach (2) as well as a GP fixed effects strategy (3). The corresponding estimates relate to somewhat different treatment effects. Thus, robustness of the results obtained by these strategies allows us to conclude on a high validity of our findings. Each of the three empirical strategies are addressed in more detail below.



First of all, the most straightforward way to correct for potential selection biases is to control for possible confounders (as far as they are observable). This approach has been chosen in all of the previously conducted German gatekeeping studies taking selection into account. We use linear regressions implying linear probability models for binary outcomes.<sup>19</sup> In a first step, we regress our outcome measures on an indicator of program participation while controlling for a large set of observable characteristics, i.e.:

$$Y_i = \alpha + \beta T_i + \gamma X_i + \varepsilon_i,$$

where  $Y_i$  represents individual  $i$ 's respective health outcome,  $T_i$  is an indicator for participation in the gatekeeping program,  $X_i$  covers our set of control variables,  $\varepsilon_i$  is the error term and  $\alpha, \beta, \gamma$  represent the estimated coefficients. Thus,  $\beta$  is supposed to provide the average treatment effect of interest. Observable controls include the characteristics presented in Table 1: We use age (third order polynomial), sex, region of residence (two-digit zip code area), a labor force indicator, patients' morbidity and an indicator on nursing care requirement as well as health care utilization, including ambulatory costs. All information refers to the year prior to the participation decision (in 2009). We assume that the controlled characteristics influence the participation decision as well as patients' future outcomes. The 17 morbidity indicators included are based on the morbidity categories commonly used for calculating the Charlson morbidity score (Charlson, 1987; Sundararajan et al., 2004). In order to account for *a priori* differences in GP quality, we also control for characteristics of the GP offices, i.e. an indicator variable for group practices, number of patients as well as offer of DMP programs (in 2009). Controlling for patient and GP characteristics in 2009 follows the logic of a difference-in-difference approach.

One may question to what extent controlling for observed characteristics may reduce potential selection biases (e.g. Riens et al., 2010). Our estimates might still suffer from biases due to unobserved characteristics influencing both the outcome measures and program participation. Thus, we do not solely rely on the assumed comprehensive specification of characteristics causing the selection biases. As an alternative strategy, we use an instrumental variable (IV) approach identifying a local average treatment effect (LATE) based on the spatial distribution of gatekeeping contracts (Imbens and Angrist, 1994), i.e.

$$\begin{aligned} (\text{first stage}) \quad T_i &= \delta + \theta R_i + \mu X_i + e_i, \\ (\text{second stage}) \quad Y_i &= \alpha_1 + \beta_1 \hat{T}_i + \gamma_1 X_i + \vartheta_i, \end{aligned}$$

---

<sup>19</sup> As an alternative to regression analysis, previous studies use matching designs. In a previous version of our study, we have conducted propensity score matching which provided robust results when compared to our regression results (own project report, results available upon request).

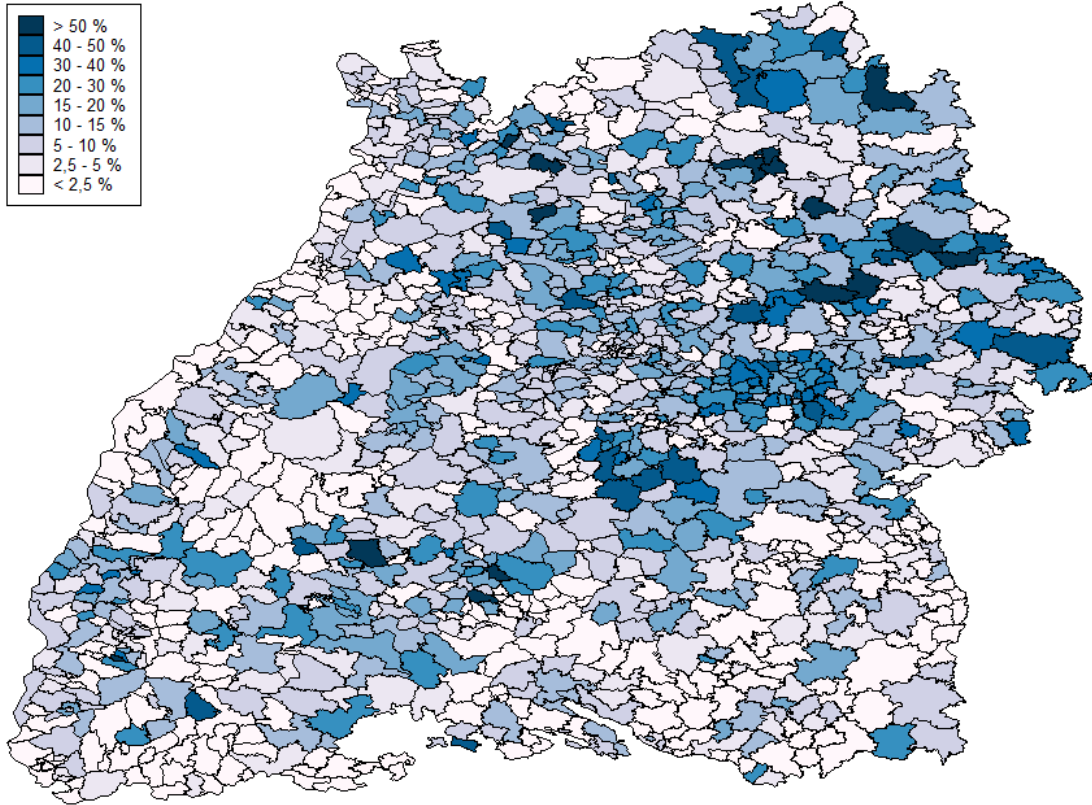
where the instrument  $R_i$  is the program participation rate within detailed (5 digit) zip-code area and  $\hat{T}_i$  is the first stage participation estimate. The same set of control variables ( $X_i$ ) is included in the first and second stage regressions.  $e_i$  ( $\vartheta_i$ ) denotes the first (second) stage error term while  $\delta, \theta$  and  $\mu$  ( $\alpha_1, \beta_1, \gamma_1$ ) are the estimated first (second) stage coefficients.

As shown in Figure 1, program participation substantially varies between zip-code areas. The logic of our instrument  $R_i$  is that we expect patients' probability to participate ( $T_i$ ) to be higher in regions with high shares of other patients opting for the contract, i.e. where more information on the existence of these contracts is available. At the same time, we assume other patients' choices to be exogenous with respect to the individual's health status and health care costs. The resulting LATE  $\beta_1$  is supposed to be representative for patients opting for gatekeeping contracts based on "what their neighbors do" within the zip-code area. The first stage estimate  $\theta$  and weak instrument check are presented in Section 5, together with the LATE estimates of  $\beta_1$  (see Table 2). We also use an alternative version of the instrument drawing on the regional participation rate *net of individuals' own GP*.<sup>20</sup>

### **Figure 1: Spatial distribution of gatekeeping contracts**

---

<sup>20</sup> Patients are more likely to participate if they live in regions with high shares of participating *practitioners*. In this respect, anecdotal evidence (based on statements of health insurance representatives) suggests that the share of participating GPs is highest in regions where representatives of the federal state GP association are most influential within the regional GP associations. Concerning our identification strategy, this would imply the assumption that the GP association's influence exerts no direct impact on patients' outcomes (i.e. beside the effects driven by the organization of gatekeeping care).



**Note:** Participation rates based on our data-set for a major health insurance company in the state of Baden-Württemberg (in 2013).

We expect that our IV identification strategy will reduce the bias of the estimated treatment effects. However, the estimated effects will still be biased if the regional provision of contracts is driven by unobserved GP characteristics also affecting other patients' participation. One might suspect that practitioners offering gatekeeping differ from non-participating GPs in characteristics also determining service quality. Therefore, we address the possibility of GP quality biases more directly. To this end, we exploit the observation of GPs treating participating as well as non-participating patients at the same time. We estimate the following GP fixed effects (FE) model:

$$(Y_{gp,i} - \overline{Y_{gp}}) = \beta_2(T_{gp,i} - \overline{T_{gp}}) + \gamma_2(X_{gp,i} - \overline{X_{gp}}) + (\omega_{gp,i} - \overline{\omega_{gp}}),$$

where  $gp$  and  $i$  index GPs and patients respectively,  $\overline{Y_{gp}}$ ,  $\overline{T_{gp}}$ ,  $\overline{X_{gp}}$  and  $\overline{\omega_{gp}}$  represent averages for GP's over patients outcome measures, participation indicators, control variables and error terms respectively. Among the estimated coefficients,  $\beta_2$  represents the treatment effect based on the within (GP) variation.

Thus, we control for within-practice quality in order to estimate a mere gatekeeping effect not confounded by GP quality. One implication of this approach is that it allows discussing

potential spill-over effects within GP offices. For example, it is possible that GPs offering the gatekeeping program adjust their behavior to treat all of their patients in the same, better coordinated way.<sup>21</sup> We will discuss the insights of the fixed effects evidence in Sections 5 and 6.

## 5 Results

Table 2 summarizes the regression results based on our optional identification strategies. In this section, we first discuss our findings from the basic regressions. Then, we describe the additional insights gained from the IV and FE estimates.

The first set of outcome indicators refers to aspects of utilization and coordination of care (see Section 3). The corresponding results from the basic regressions in column (1) suggest that gatekeeping principally does what it is supposed to do: The number of GPs visited within one year is significantly reduced (about 0.16) for patients participating in the program. Also, the number of specialist visits *with* referral is significantly higher (0.49) while the number of specialist visits *without* reported referral is lower (-0.43). Thus, the reported *share* of specialist visits with referral is higher for participants (43 % vs. 31 % for non-participants, not directly shown in Table 2). At the same time, *ceteris paribus* the total number of specialist treatment cases is not significantly different for participants and non-participants in the gatekeeping program.

**Table 2: Estimated impacts of gatekeeping based on optional empirical strategies**

	Sample average	(1) Basic regressions	(2) Instrumental variable	(3) GP fixed effects
<i>Coordination/continuity of health care</i>				
Number of GPs	1.55	-0.16 *** (0.00)	-0.30 *** (0.01)	-0,15 *** (0,01)
Specialist visits with referral	1.35	0.49 *** (0.01)	0.51 *** (0.03)	0,65 *** (0,02)
Specialist visits without referral	2.70	-0.43 *** (0.02)	-0.55 *** (0.05)	-0,34 *** (0,02)
<i>Participation in DMPs of newly diagnosed...</i>				
... DMP asthma	0.04	0.02 ** (0.01)	0.04 (0.03)	0,01 (0,02)
... DMP COPD	0.05	0.02 * (0.01)	0.04 * (0.03)	0,01 (0,02)
... DMP diabetes type II	0.29	0.12 ***	0.10 **	0,09 ***

<sup>21</sup> In the extreme case, we will always obtain zero effects of the gatekeeping program and this might be considered as a source of bias of our identification strategy. However, this is not what we find in the empirical analysis.

... DMP CHD	0.08	(0.02) 0.02 ** (0.01)	(0.04) 0.05 ** (0.02)	(0,03) 0,03** (0,02)
<i>Health care quality</i>				
Influenza vaccination	0.35	0.07 *** (0.00)	0.05 *** (0.13)	0,08 *** (0,01)
General health check-up	0.27	0.02 *** (0.00)	0.02 *** (0.01)	0,02 *** (0,00)
ACSH	0.02	-0.00 ** (0.00)	-0.01 ** (0.00)	-0,00 *** (0,00)
<i>Cost indicator</i>				
Log ambulatory costs <sup>A</sup>	555.11	0.15 *** (0.00)	0.09 *** (0.15)	0,23 *** (0,01)
First stage: regional participation rate (instrument)			1.16 *** (0.01)	
First stage partial F-statistic			17,895.6	
Sample size		281.635		

**Note:** The first column of figures refers to average numbers (<sup>A</sup>average not in log for cost indicator). All other numbers correspond to estimated coefficients (with standard errors in parentheses), unless otherwise stated. \* Significant at the 10 % level, \*\* 5 % level, \*\*\* 1 % level.

**Source:** Own estimations based on data for a major health insurance company in the German state of Baden-Württemberg (in 2013, control variables relate to the year before introduction of the gatekeeping program in 2009).

In line with a better coordination of care, the probability to enroll in DMPs is significantly higher for participants in the gatekeeping program as compared to the control group. The impact amounts to a two percentage point increase for asthma, COPD and CHD. For diabetes type II, the impact amounts to 12 percentage points. These are substantial increases, given that the enrolled proportion of recently diagnosed patients is generally rather low (varying between 4 % for asthma and 27 % for type II diabetes in the control group).

The middle panel of Table 2 presents the estimates for the health care quality indicators. Our findings suggest that the use of preventive measures is significantly higher for gatekeeping participants than for non-participants. The impact amounts to seven percentage points for influenza vaccinations and to two percentage points for health check-ups. Again, these are substantial effects: 34 % of the elderly patients receive the flu shot in the control group and about 27 % get check-ups. Similarly, concerning the avoidance of unnecessary hospitalizations, the probability of ACSH is reduced by about 0.2 percentage points for gatekeeping participants. This corresponds to about 10 %, referring to an average of 2 % in the control group.

The coordination and quality effects seem to go along with significantly higher costs of ambulatory care. The respective estimate suggests that annual ambulatory health care costs are about 16 % higher for gatekeeping patients as compared to the control group.

The optional IV regression results are presented in column (2) of Table 2. Information on the instrument is provided at the bottom of Table 2. The first stage coefficient of 1.16 implies that an increase in the regional participation rate of about 10 percentage points yields a 12 percentage points increase in the individual participation probability. The very high partial F-statistics suggests that the instrument does not suffer from a weak instrument problem (cf. Staiger and Stock, 1997; Stock, Wright, and Yogo, 2002).

In general, while being robust, the point estimates of the gatekeeping effects based on IV regressions tend to be higher as compared to the basic regression results in column (1). The estimates are consistently higher when looking at health care utilization. We assume that this is due to the specific local average treatment effect estimated by IV. Compliers in terms of our instrument (“opting for the contract when their neighbors do so”) might be a group of patients also acting more compliant with respect to the gatekeeping contracts’ conditions (e.g. requiring to see the gatekeeping GP for referrals).

As an alternative to the considered instrument, we have used regional participation rates *net of patients visiting the respective individual’s GP* (see Appendix, Table A1). In this case, we still have a strong instrument (with a first stage coefficient of 0.5) and fairly robust second stage results.

Column (3) of Table 2 presents the respective estimates from the GP fixed effects regressions. Again, the estimates are fairly robust as compared to the previous regression results. The coefficient on specialist visits with (without) referral is somewhat higher (lower) as compared to columns (1) and (2). This means that within GP offices, gatekeeping participants and non-participants are treated in (more) differential ways regarding this aspect of the coordination of care. However, the point estimates of the quality indicators are strongly robust among regressions. Even within GP offices (i.e. net of GP quality), we observe significant and substantial quality impacts of patients’ participation in the gatekeeping program. One interpretation of the robust finding is that there seem to be no positive quality spill-over effects from GPs participation in the gatekeeping program (also see discussion in the next section).

## **6 Discussion**

GPs role as gatekeepers is discussed as a means to increase efficiency in health care services in many countries (EXPH, 2014). As such, the primary objective of the German gatekeeping schemes is to increase overall quality of care while ensuring cost effectiveness (BT-Drs. 16/3100, 2006; BT-Drs. 16/7576, 2008).

The results of our evaluation study suggest that the German gatekeeping contract yields a somewhat higher coordination of care, improved quality (regarding prevention and avoidance of hospitalization) but also higher ambulatory costs. Our empirical results are largely robust among the identification designs. Also, the sizes of the quality effects are substantial: For influenza vaccinations among elderly patients, the average effect across specifications is about 0.07, corresponding to a 20 % increase in the population mean. Similarly, gatekeeping patients more often obtain general health check-ups with a mean effect size of 0.02 (7 % of the sample mean). Also, patients under gatekeeping suffer less often from avoidable hospitalizations (average over all estimation results of -0.003 or -17 % of the sample mean).

The findings are in line with findings from the previous literature on gatekeeping in Germany. Particularly, the most recent and comprehensive study by Gerlach and Szecsenyi (2016) reports similar findings for avoidable hospitalizations (about 1 percentage point). For influenza vaccinations among the elderly, Gerlach and Szecsenyi (2016) find an effect of 4.70 % (roughly comparable to our estimate of 6.9 %). The study is based on claims data of another major insurance company in Germany. Due to the similarity of results, we infer that our gatekeeping effects may also apply to other health insurance companies' populations.

The overall positive quality effects also provide guidance toward the financial incentives potentially interacting with the gatekeeping effects. As detailed in Section 3, GPs operate under two distinct financial schemes: While they are remunerated mainly on a fee-for-service scheme for the non-gatekeeping patients, they obtain higher capitation payments for the gatekeeping patients (see Section 3 for details). For health check-ups, GPs are remunerated equally for both groups of patients; financial incentives may thus be considered of limited concern for identifying the respective gatekeeping effect. On the other hand, one may expect a poor performance of the (gatekeeping) capitation system on overall health care quality: Profit-maximizing GPs may be incentivized to limit treatment quality under the capitation scheme. However, we still observe a substantial beneficial gatekeeping impact on avoided hospitalizations. Thus, the capitation element of the gatekeeping contracts seems not to (entirely) counteract the potentially improved quality through better coordination. The overall positive quality estimates raise confidence in the gatekeeping services actually improving aspects of care quality.

A further insight is gained when considering the results from the GP fixed effects specifications. One may expect that participating GPs that are becoming gatekeepers for some of their patients generally adjust their treatment behavior. If these GPs offered the same treatment to all of their patients, the fixed effects estimates would be biased towards zero. However, this does not seem to be the case as the FE point estimates tend to be

higher than the estimates from the basic regressions as well as the IV regressions. We conclude that there are no positive spill-over effects within the GP office.

Overall, our results suggest that the gatekeeping contract may yield the quality effects envisioned by policy-makers supporting a strong primary care system. However, the limited number of quality indicators we observe only depict selected elements of the overall care environment. Also, finding that the gatekeeping yields higher ambulatory costs is only a part of the story. In order to come to a conclusion on the overall benefits of the system, a more extensive cost-benefit analysis is required. Similarly, the time horizon of our analysis is rather limited. One may expect that quality outcomes strengthen over time (Gerlach and Szecsenyi, 2016). To this end, our current analysis of indicators relating to health coordination, process- and outcome-related quality and ambulatory costs for the years 2009-2013 provides first insights and may be viewed as being a first step towards a more comprehensive longer-run analysis.



## References

- Abbasi, K. (2014). The demise of gatekeeping in primary care. *Journal of the Royal Society of Medicine*, 107(11), 427, DOI: 10.1177/0141076814557606.
- Allard, M., Jelovac, I., Leger, P.T. (2011). Treatment and referral decisions under different physician payment mechanisms. *Journal of Health Economics*, 30(5), 880-893, DOI: 10.1016/j.jhealeco.2011.05.016.
- AQUA, 2009. *QiSA Band E1. Prävention. Qualitätsindikatoren für die Vermeidung von Krankheiten*, AQUA Institut für Angewandte Qualitätsförderung und Forschung im Gesundheitswesen, AOK Bundesverband, Berlin.
- AQUA, 2013. *Evaluation von fünf Ersatzkassen-Hausarztverträgen auf Grundlage des §73b SGB V*. AQUA Institut Göttingen, <[https://www.aqua-institut.de/aqua/upload/CONTENT/Projekte/2013-09-16\\_HzV-Evaluation\\_vdek.pdf](https://www.aqua-institut.de/aqua/upload/CONTENT/Projekte/2013-09-16_HzV-Evaluation_vdek.pdf)>, last accessed 5 August 2016.
- Blomqvist, Å., Leger, P.T. (2005). Information asymmetry, health insurance and the decision to hospitalize. *Journal of Health Economics*, 24(4), 775-793.
- Böcken, J. (2008). Hausarztmodelle im Spannungsfeld zwischen ordnungspolitischem Anspruch und Versorgungsrealität. In Böcken, J., Braun, B., Amhof, R. (editors), *Gesundheitsmonitor 2008*, Verlag Bertelsmann Stiftung.
- Brekke, K. R., Nuscheler, R., Straume, O.R. (2007). Gatekeeping in Health Care, *Journal of Health Economics*, 26(1), pp. 149-170, DOI: 10.1016/j.jhealeco.2006.04.004.
- BT-Drs. 16/3100 (2006). *Drucksache des Deutschen Bundestages 16/3100 vom 24.10.2006*. Entwurf eines Gesetzes zur Stärkung des Wettbewerbs in der gesetzlichen Krankenversicherung (GKV-Wettbewerbsstärkungsgesetz – GKV-WSG).
- BT-Drs. 16/7576 (2008). *Drucksache des Deutschen Bundestages 16/7576 vom 02.01.2008*. Antwort der Bundesregierung auf die Kleine Anfrage der Abgeordneten Frank Spieth, Klaus Ernst, Dr. Martina Bunge, und weiterer Abgeordneter.
- Charlson, M.E., Pompei, P., Ales, K.L., MacKenzie, C.R. (1987). A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation, *Journal of Chronic Diseases*, 40(5), 373-83, DOI: 10.1016/0021-9681(87)90171-8.
- Destatis (2013). *Fachserie 13, Reihe 1.1. Sozialleistungen. Angaben zur Krankenversicherung (Ergebnisse des Mikrozensus)*. Statistisches Bundesamt.
- Docteur E., Oxley H. (2003). *Health-care systems: Lessons from the reform experience*. Organization for Economic Cooperation and Development (OECD), Paris.
- Donabedian, A. (1988). The Quality of Care – How Can It Be Assessed, *Journal of the American Medical Association (JAMA)*, 260(12), 1743-1748, DOI: 10.1001/jama.260.12.1743.
- EXPH (2014). *Report on Definition of a frame of reference in relation to primary care with a special emphasis on financing systems and referral systems*. Expert Panel on effective ways of investing in Health, European Union.
- Freund, T., Szecsenyi, J., Ose, D. (2010). Charakteristika von eingeschriebenen Versicherten eines flächendeckenden Vertrages zur hausarztzentrierten Versorgung, *Medizinische Klinik*, 105(11), 808-811.
- Garrido, M. V., Zentner, A., Busse, R. (2011). The effects of gatekeeping: A systematic review of the literature. *Scandinavian Journal of Primary Health Care*, 29(1), 28-38, DOI: 10.3109/02813432.2010.537015.

- Gaynor, M., Town, R.J. (2011). *Competition in Health Care Markets*. Working Paper 17208, National Bureau of Economic Research (NBER).
- Gerdtham, U., Jönsson, B., Macfarlan, M., Oxley, H. (1998) . The determinants of health expenditure in the OECD countries: a pooled data analysis. In Zweifel, P. (editor), *Health, the medical profession, and regulation*, Kluwer Academic Publishers, 113-134.
- Gerdtham, U.-G., Jönsson, B. (2000). International Comparisons of Health Expenditure. In Culyer, A.J., Newhouse, J.P. (editors), *Handbook of Health Economics*, Volume 1A.
- Gerlach, F., Szecsenyi, J. (2013a). *Abschlussbericht zur Evaluation der Hausarztzentrierten Versorgung (HzV) nach § 73b SGB V in Baden-Württemberg (2010-2011)*, <[https://www.aok-gesundheitspartner.de/imperia/md/gpp/bw/arztundpraxis/hzv/bw\\_hzv\\_eval\\_2008\\_2010.pdf](https://www.aok-gesundheitspartner.de/imperia/md/gpp/bw/arztundpraxis/hzv/bw_hzv_eval_2008_2010.pdf)>, last accessed 5 August 2016.
- Gerlach, F., Szecsenyi, J. (2013b). Hausarztzentrierte Versorgung in Baden-Württemberg – Konzept und Ergebnisse der kontrollierten Begleitevaluation, *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen*, 107, 365-371.
- Gerlach, F., Szecsenyi, J. (2014). *Evaluation der Hausarztzentrierten Versorgung (HzV) nach §73b SGB V in Baden-Württemberg (2013-2016)*. Ergebnisbericht, <[https://www.aok-gesundheitspartner.de/imperia/md/gpp/bw/arztundpraxis/hzv/bw\\_hzv\\_eval\\_2013-2014.pdf](https://www.aok-gesundheitspartner.de/imperia/md/gpp/bw/arztundpraxis/hzv/bw_hzv_eval_2013-2014.pdf)>, last accessed 5 August 2016.
- Gerlach, F., Szecsenyi, J. (2016). *Evaluation der Hausarztzentrierten Versorgung (HZV) in Baden-Württemberg, Zusammenfassung der Ergebnisse – Ausgabe 2016*, Frankfurt/Heidelberg (project report), <[http://neueversorgung.de/images/PDF/Evaluation\\_2016/HZV\\_Evaluation\\_Broschuere.pdf](http://neueversorgung.de/images/PDF/Evaluation_2016/HZV_Evaluation_Broschuere.pdf)>, last accessed 5 August 2016.
- Godager, G., Iversen, T., Ma, C.A. (2015). Competition, gatekeeping, and health care access, *Journal of Health Economics*, 39, 159-170, DOI: 10.1016/j.jhealeco.2014.11.005.
- Gross, P.A., Hermogenes, A.W., Sacks, H.S., Lau, J, Levandowski, R. A. (1995). The efficacy of influenza vaccine in elderly persons. A meta-analysis and review of the literature. *Annals of Internal Medicine*, 123(7), 518-527.
- Hansen, J., Groenewegen, P.P., Kringos, D.S. (2015). Strength of primary care and health outcomes for chronically ill people: results from 27 EU Member States, *Health Affairs*, 34(9), 1531-1537.
- Harrison, M. J., Dusheiko, M., Sutton, M., Gravelle, H., Doran, T., Roland, M. (2014). Effect of a National Primary Care Pay For Performance Scheme on Emergency Hospital Admissions for Ambulatory Care Sensitive Conditions: Controlled Longitudinal Study, *British Medical Journal*, 349, g6423, DOI: 10.1136/bmj.g6423.
- Hausärzteverband (2015). *Hausarztzentrierte Versorgung*, Sonderbeilage in Der Hausarzt / Der Allgemeinarzt, Februar/März 2015.
- Hausärzteverband (2016). [*Gatekeeping contracts, in German*], <<http://www.hausaerzteverband.de/cms/Hausarztvertraege.988.0.html>>, last accessed 5 August 2016.
- Hjortdahl, P., Borchgrevink, C.F. (1991). Continuity of care: Influence of general practitioners' knowledge about their patients on use of resources in consultations. *British Medical Journal*, 303, 1181-1184, DOI: 10.1136/bmj.303.6811.1181.
- Hoebel, J., Richter, M., Lampert, T. (2013). Social status and participation in health checks in men and women in Germany – results from the German Health Update (GEDA), 2009 and 2010. *Deutsches Ärzteblatt International*, 110(41), 679-85.

- Hofmann, S. Mühlenweg, A.M. (2015). *Tür auf für den Gatekeeper? Konzeption eines Indikatoren-Sets für ein routinemäßiges Prozess-, Ergebnis- und Wirtschaftlichkeits-Monitoring der hausarztzentrierten Versorgung*, WifOR, mimeo.
- Imbens, G., Angrist, J. (1994). Identification and estimation of local average treatment effects, *Econometrica*, 62(2), 467-476, DOI: 10.2307/2951620.
- Imbens, G. (2015). Matching Methods in Practice. Three Examples. *Journal of Human Resources*, 50(2), 373-419, DOI: 10.3368/jhr.50.2.373.
- Krankenkasseninfo (2016): *Mitglieder und Versicherte je Krankenkasse*, <http://www.krankenkasseninfo.de/zahlen-fakten/mitgliederzahlen/>, last accessed 5 August 2016.
- Kringos et al. (2010). The European Primary Care Monitor: Structure, Process and Outcome Indicators. *BMC Family Practice*, 11(1), DOI: 10.1186/1471-2296-11-81.
- Kringos, D.S., Boerma, W., van der Zee, J., Groenewegen, P. (2013). Europe's strong primary care systems are linked to better population health but also to higher health spending. *Health Affairs*, 32(4), 686-694, DOI: 10.1377/hlthaff.2012.1242.
- Kürschner, N., Weidmann, W., Müters, S. (2011). Wer wählt den Hausarzt zum „Gatekeeper“? Eine Anwendung des Verhaltensmodells von Andersen zur Beschreibung der Teilnahme an Hausarztmodellen in Deutschland, *Bundesgesundheitsblatt*, 54, 221-227.
- Mariñoso, B.G., Jelovac, I. (2003). GP's payment contracts and their referral practice. *Journal of Health Economics*, 22, 617-635, DOI: 10.1016/S0167-6296(03)00008-0.
- Mossialos, E., Wenzl, M., Osborn, R., Sarnak, D. (2016). *2015 International Profiles of Health Care Systems*, The Commonwealth Fund, pub. 1857, January 2016.
- McGuire, T. (2000). Physician Agency. In Culyer, A.J., Newhouse, J.P. (editors), *Handbook of Health Economics*, Volume 1A.
- Nichol, K.L., Wuorenma, J., von Sternberg, T. (1998). Benefits of influenza vaccination for low-, intermediate-, and high-risk senior citizens. *Archives of Internal Medicine*, 158(16), 1769-1776, DOI: 10.1001/archinte.158.16.1769.
- Ose, D., Broge, B., Riens, B., Szecsenyi, J. (2008). Mit Überweisung vom Hausarzt zum Spezialisten – Haben Verträge zur Hausarztzentrierten Versorgung (HzV) einen Einfluss? *Zeitschrift für Allgemeinmedizin*, 84, 321-326.
- Riens, B., Broge, B., Kaufmann-Kolle, P., Pöhlmann, B., Grün, B., Ose, D., Szecsenyi, J. (2010). Bildung einer Kontrollgruppe mithilfe von Matched-Pairs auf Basis von GKV-Routinedaten zur prospektiven Evaluation von Einschreibemodellen, *Das Gesundheitswesen*, 72, 363-370.
- Scott, A. (2000). Economics of general practice. In Culyer, A.J., Newhouse, J.P. (editors), *Handbook of Health Economics*, Volume 1.
- Staiger, D., Stock, J.H. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica*, 65(3), 557-586, DOI: 10.2307/2171753.
- Steiner, M., Riedel, W., Maetzel, J. (2013). *Wissenschaftliche Begleitung des hausärztlichen Qualitäts- und Kooperationsmodells Südbaden*, Abschlussbericht, Prognos AG, <[http://www.aokgesundheitspartner.de/imperia/md/gpp/bw/arztundpraxis/arztnetze/bw\\_arztnetz\\_ze\\_abschlussbericht.pdf](http://www.aokgesundheitspartner.de/imperia/md/gpp/bw/arztundpraxis/arztnetze/bw_arztnetz_ze_abschlussbericht.pdf)>, last accessed 5 August 2016.
- Stock, James H., Jonathan H. Wright, and Motohiro Yogo. 2002. Survey of Weak Instruments and Weak Identification in Generalized Method of Moments. *Journal of Business and Economic Statistics*, 20(4), 518-529, DOI: 10.2307/2171753.

- Sundararajan, V., Henderson, T., Perry, C., Muggivan, A., Quan, H., Ghali, W.A. (2004). New ICD-10 version of the Charlson Comorbidity Index predicted in-hospital morbidity. *Journal of Clinical Epidemiology*, 57, 1288-1294.
- Sundmacher, L., Kopetsch, T. (2015). The impact of office based care on hospitalizations for ambulatory care sensitive conditions. *European Journal of Health Economics*, 16, 365-375, DOI: 10.1007/s10198-014-0578-4.
- Sundmacher, L., Fischbach, D., Schuettig, W., Naumann, C., Augustin, U., Faisst, C. (2015). Which hospitalisations are ambulatory care-sensitive, to what degree, and how could the rates be reduced? Results of a group consensus study in Germany. *Health Policy*, 119(11), 1415-1423, DOI: 10.1016/j.healthpol.2015.08.007.
- Vedsted, P., Olesen, F (2011). Are the serious problems in cancer survival partly rooted in gatekeeper principles? An ecologic study. *British Journal of General Practice*, 61(589), e508-e512, DOI: 10.3399/bjgp11X588484.
- Weiss, L.J., Blustein, J. (1996). Faithful patients: The effect of long-term physician-patient relationships on the costs and use of health care by older Americans. *American Journal of Public Health*, 86, 1742-1747, DOI: 10.2105/AJPH.86.12.1742.
- Wilde, J.A., McMillan, J.A., Serwint, J., Butta, J., O'Riordan, M.A., Steinhoff M.C. (1999). Effectiveness of influenza vaccine in health care professionals: a randomised trial. *Journal of the American Medical Association*, 281(10), 908-913, DOI: 10.1001/jama.281.10.908.

## Appendix

**Table A1: Robustness of estimated impacts of gatekeeping, alternative IV results**

	Sample average	(1) Basic regressions	(2) Instrumental variable	(3) Instrumental variable (II)
<i>Coordination/continuity of health care</i>				
Number of GPs	1.55	-0.16 *** (0.00)	-0.30 *** (0.01)	-0.55 *** (0.03)
Specialist visits with referral	1.35	0.49 *** (0.01)	0.51 *** (0.03)	0.65 *** (0.08)
Specialist visits without referral	2.70	-0.43 *** (0.02)	-0.55 *** (0.05)	-0.54 *** (0.12)
<i>Participation in DMPs of newly diagnosed...</i>				
... DMP asthma	0.04	0.02 ** (0.01)	0.04 (0.03)	0.06 (0.07)
... DMP COPD	0.05	0.02 * (0.01)	0.04 * (0.03)	0.04 (0.07)
... DMP diabetes type II	0.29	0.12 *** (0.02)	0.10 ** (0.04)	0.10 (0.10)
... DMP CHD	0.08	0.02 ** (0.01)	0.05 ** (0.02)	0.10 (0.07)
<i>Health care quality</i>				
Influenza vaccination	0.35	0.07 *** (0.00)	0.05 *** (0.13)	-0.00 (0.03)
General health check-up	0.27	0.02 *** (0.00)	0.02 *** (0.01)	0.02 (0.02)
ACSH	0.02	-0.00 ** (0.00)	-0.01 ** (0.00)	-0.01 *** (0.01)
<i>Cost indicator</i>				
Log ambulatory costs <sup>A</sup>	555.11	0.15 *** (0.00)	0.09 *** (0.15)	0.07 *** (0.04)
First stage: regional participation rate (instrument)			1.16 *** (0.01)	0.50 *** (0.01)
First stage partial F-statistic			17,895.6	3,098.17
Sample size		281.635		

**Note:** The alternative instruments are regional participation rates (column 2, also shown in Table 2) and regional participation rates *net of own GP office* (instrumental variable II, column 3). The first column of figures refers to average numbers (<sup>A</sup>average not in log for cost indicator). All other numbers correspond to estimated coefficients (with standard errors in parentheses), unless otherwise stated. \* Significant at the 10% level, \*\* 5% level, \*\*\* 1% level.

**Source:** Own estimations based on data for a major health insurance company in the German state of Baden-Württemberg (in 2013, control variables relate to the year before introduction of the gatekeeping program in 2009).



## CINCH working paper series

- 1 Halla, Martin and Martina Zweimüller. **Parental Responses to Early Human Capital Shocks:** Evidence from the Chernobyl Accident. CINCH 2014.
- 2 Aparicio, Ainhoa and Libertad González. **Newborn Health and the Business Cycle:** Is it Good to be born in Bad Times? CINCH 2014.
- 3 Robinson, Joshua J. **Sound Body, Sound Mind?:** Asymmetric and Symmetric Fetal Growth Restriction and Human Capital Development. CINCH 2014.
- 4 Bhalotra, Sonia, Martin Karlsson and Therese Nilsson. **Life Expectancy and Mother-Baby Interventions:** Evidence from A Historical Trial. CINCH 2014.
- 5 Goebel, Jan, Christian Krekel, Tim Tiefenbach and Nicolas R. Ziebarth. **Natural Disaster, Environmental Concerns, Well-Being and Policy Action:** The Case of Fukushima. CINCH 2014.
- 6 Avdic, Daniel, **A matter of life and death? Hospital Distance and Quality of Care:** Evidence from Emergency Hospital Closures and Myocardial Infarctions. CINCH 2015.
- 7 Costa-Font, Joan, Martin Karlsson and Henning Øien. **Informal Care and the Great Recession.** CINCH 2015.
- 8 Titus J. Galama and Hans van Kippersluis. **A Theory of Education and Health.** CINCH 2015.
- 9 Dahmann, Sarah. **How Does Education Improve Cognitive Skills?:** Instructional Time versus Timing of Instruction. CINCH 2015.
- 10 Dahmann, Sarah and Silke Anger. **The Impact of Education on Personality:** Evidence from a German High School Reform. CINCH 2015.
- 11 Carbone, Jared C. and Snorre Kverndokk. **Individual Investments in Education and Health.** CINCH 2015.
- 12 Zilic, Ivan. **Effect of forced displacement on health.** CINCH 2015.

- 13 De la Mata, Dolores and Carlos Felipe Gaviria. **Losing Health Insurance When Young:** Impacts on Usage of Medical Services and Health. CINCH 2015.
- 14 Tequame, Miron and Nyasha Tirivayi. **Higher education and fertility:** Evidence from a natural experiment in Ethiopia. CINCH 2015.
- 15 Aoki, Yu and Lualhati Santiago. **Fertility, Health and Education of UK Immigrants:** The Role of English Language Skills. CINCH 2015.
- 16 Rawlings, Samantha B., **Parental education and child health:** Evidence from an education reform in China. CINCH 2015.
- 17 Kamhöfer, Daniel A., Hendrik Schmitz and Matthias Westphal. **Heterogeneity in Marginal Non-monetary Returns to Higher Education.** CINCH 2015.
- 18 Ardila Brenøe, Anne and Ramona Molitor. **Birth Order and Health of Newborns:** What Can We Learn from Danish Registry Data? CINCH 2015.
- 19 Rossi, Pauline. **Strategic Choices in Polygamous Households:** Theory and Evidence from Senegal. CINCH 2016.
- 20 Clarke, Damian and Hanna Mühlrad. **The Impact of Abortion Legalization on Fertility and Maternal Mortality:** New Evidence from Mexico. CINCH 2016.
- 21 Jones, Lauren E. and Nicolas R. Ziebarth. **US Child Safety Seat Laws:** Are they Effective, and Who Complies? CINCH 2016.
- 22 Koppensteiner, Martin Foureaux and Jesse Matheson. **Access to Education and Teenage Pregnancy.** CINCH 2016.
- 23 Hofmann, Sarah M. and Andrea M. Mühlenweg. **Gatekeeping in German Primary Health Care –** Impacts on Coordination of Care, Quality Indicators and Ambulatory Costs. CINCH 2016.





