Population Change and the Regional Distribution of Physicians

Michael Kuhn, Carsten Ochsen

PII: S2212-828X(18)30052-5
DOI: https://doi.org/10.1016/j.jeoa.2019.100197
Article Number: 100197
Reference: JEOA 100197

To appear in: *The Journal of the Economics of Ageing*

Please cite this article as: M. Kuhn, C. Ochsen, Population Change and the Regional Distribution of Physicians, *The Journal of the Economics of Ageing* (2019), doi: https://doi.org/10.1016/j.jeoa.2019.100197

Population Change and the Regional Distribution of Physicians*

Michael Kuhn† Carsten Ochsen‡

February 25, 2019

Abstract

Based on an intertermporal entry model of the physician market, we analyze how the supply of office-based physicians depends on regional character and on the age-structure of the local population as determinants of the current profitability of physician services, on local population change as a predictor of future demand, and on the extent of equilibrium adjustment within local markets. Using German regional data, we find that the number of general practitioners (GPs) per capita is positively related to the share of the population 60 and above within metropolitan areas, but negatively within rural areas. Future changes in list size have an impact on the current supply of GPs, suggesting limitations to equilibrium adjustment especially in regions with excess supply. Overall, population change should have raised the profitability of GP services over the period 1997–2008. The falling supply of GPs, especially in rural regions, then implies an increase in reservation income.

Keywords: age structure, entry equilibrium, inequality in health care, panel data, regional physician supply, population ageing.

JEL classification: C23, I11, J10, J44, R23.

---

*We are grateful to two anonymous reviewers for their constructive comments. All responsibility for remaining errors lies with us.

†Wittgenstein Centre (IIASA, VID/ÖAW, WU), Vienna Institute of Demography, International Institute for Applied Systems Analysis, Austria; Correspondence: michael.kuhn@oeaw.ac.at.

‡University of Applied Labour Studies and University of Rostock.

1 Introduction

Population ageing is widely expected to come with an increased per-capita demand for ambulatory physician services. According to a popular argument regions with high population shares of elderly persons should then be particularly profitable for physician practice and should therefore exhibit a high number of physicians per capita.[1] This view contrasts with recent evidence from Germany according to which rural regions with large elderly populations are in danger of being under-doctored. In these regions, ageing may lead to a widening gap between the demand for health care and its supply, a situation potentially warranting policy intervention (Advisory Council 2014). Similar concerns have been raised in other countries, e.g. in the US (Cooper et al. 2002, Colwill et al 2008).

The question under what conditions and in which geographical context the health care needs of an ageing population can be adequately covered adds a new dimension to a long-standing debate on regional imbalances in the distribution of physicians.[2] In Germany, this debate has seen a turn of the tide during the late 1990s and early 2000s. Set against the general view of ubiquitous over-supply, at least in western Germany, concerns have been increasing over the last decade about physician shortages in rural regions, most notably but not exclusively in eastern Germany. While in the time span 1995–2009, the total number of office-based physicians has continued to increase by some 27 percent, there has been considerable geographical variation in this trend with a large number of regions losing physicians. At the same time, the number of general practitioners (GPs) has been falling by some 7 percent on average, the decline being much more pronounced in rural regions, where GPs are often the sole providers of medical services.

[1] Frequently, the number of physicians per capita is referred to as "physician density". We refrain from this terminology in order to avoid confusion with the measure of "population density", which we will also be referring to.

[2] Recent contributions include Ono et al. (2014) for the OECD; Fülöp et al. (2008) and Gächter et al. (2012) for Austria; Hann and Gravelle (2004), Elliott et al. (2006) and Goddard et al. (2010) for England (and Wales); Fülöp et al. (2008), Klose and Rehbein (2011), Scholz et al. (2015), Sundmacher and Ozegowski (2016) and Vogt (2016) for Germany; Iversen and Kopperud (2005) for Norway; Correia and Veiga (2010) for Portugal; Nocera and Wanzenried (2008) for Switzerland; Cooper et al. (2002) and Rosenthal et al. (2005) for the US.

Indeed, the concern about an increasingly unequal geographical distribution in the access to primary health care and about the ensuing health outcomes such as avoidable cancer deaths (Sundmacher and Busse 2011) or avoidable hospitalizations (Sundmacher and Kopetsch 2015) has been at the heart of recent health care reforms (Ozegowski and Sundmacher 2012, 2014; Advisory Council 2014). Ono et al. (2014) summarize similar policy concerns for a number of OECD countries.

It is the aim of this study to identify some of the relationships underlying the development of the geographic distribution of physicians and, in particular, to understand the role of regional population change. We construct an overlapping-generations model in which physicians commit to a practice location when young and then provide services to the local population over their working lives. Determining within entry equilibrium the number of physicians per capita within a region, we show how the regional supply of physicians depends on the stream of current and anticipated practice income and, thus, how it depends on the current and future demographic make-up, as measured by the age structure and size of the population. From the entry equilibrium condition we then derive a structural equation, allowing us to formulate and test in a theory-grounded way a number of hypotheses about the determinants of physician supply at the regional level. Our analysis points at three particularly relevant issues when it comes to assessing the impact on physician supply of regional population change. Each of these suggests that the view that regions with large elderly populations are attractive for physicians may be too simplistic.

First, the profitability of offering services to different age groups of the population does not only vary with the demand by these groups but also with the profit margin earned on each patient. If reimbursement rates are only imperfectly adjusted to differences in treatment costs and if the treatment of older patients is more costly, then it is no longer clear that a larger share of elderly patients automatically raises profit. Hence, the design of the reimbursement system is bound to matter. Furthermore, it is likely that the sensitivity of treatment costs and demand with respect to age varies systematically with the regional circumstances. For instance, long travelling distances and poor availability of public transport within rural as opposed to urban settings may render the treatment of elderly and frail patients more costly, both as more home visits are required and as rendering such visits is more costly for the physician. Thus, we expect the relative profitability of different age groups to vary with the position of a region on the urban–rural spectrum as well as with the extent to which the reimbursement system is accounting for such variation.

Second, as physicians typically set up their location for a long period – in many instances for the remainder of their working lives – their location choice should not only reflect current but also future profit opportunities. With expectations about the future demand for services being linked to population change, regions that experience population growth may attract physicians in excess of the number that would be supported by the current size of the population. Conversely, in regions with a shrinking population the number of physicians may fall short of the number that would be supported by the current population.

Third, the extent to which the future income stream bears on the current supply of physicians depends on how far the process of exit and entry leads to an adjustment in the number of physicians towards its equilibrium value. We show that under instantaneous adjustment to equilibrium, the number of physicians per capita is determined exclusively by current income, whereas future demographic development does not bear on the current supply of physicians. This is because changes in income are fully offset by compensating changes in the number of rivalling physicians. Thus, regardless of the future size and profitability of the regional market, a physician's remaining life-cycle income corresponds to the (reservation) value that is necessary for physicians to locate within that region. In contrast, if the local supply of physicians does not fully adjust towards its equilibrium value, then the current number of physicians per capita depends on future demand and income as long as the market is expected to be out of equilibrium. Reasons for a lack of adjustment towards equilibrium may be regulatory entry restrictions and the presence of sunk costs, leading to entry and exit barriers, respectively. Given the presence of entry regulation at regional level in Germany, we seek to establish empirically whether or not this leads to a substantive lack in equilibrium adjustment in local physician markets.

We use panel data for Germany at the regional level to examine empirically the relationship between local physician supply and its demographic and geographic determinants. Using annual data for 409 districts for the period 1995–2009, we analyze how the population share of individuals aged 60 and above (60plus) and the share of individuals aged 19 and below (19minus) affect the local supply of physicians. In our estimations, we focus on GPs as a large and homogeneous sub-group of office-based physicians, with at least some GPs operating in all regions. Moreover, GPs typically provide services to a local population, implying that inter-regional patient flows are of little relevance. We control for regional characteristics which, according to the literature, are relevant for the geographical distribution of physician practices.

Using a fixed-effects panel data estimator, we analyze how the number of GPs per capita depends on the age structure and growth of the population as well as on the presence of regulatory entry ceilings. Our results show that the impact of the age structure on the supply of physicians depends on whether or not a region is urban. Most prominently, the share of the population 60plus has a positive impact on GP supply within metropolitan areas but a negative impact within rural areas. Expectations about future demography, as measured by the growth rate of the list size, i.e. the number of residents served by each physician within the region, in the year subsequent to the current year of observation, have an economically small impact on the current number of GPs per capita. While this hints at some imperfections in equilibrium adjustment in general, we also show that these effects are stronger in regions for which an "excess supply" has been established and which are therefore likely to be subject to entry ceilings. We also find confirming evidence for the notion that greater population density (across space) tends to induce stronger competition, leading ceteris paribus to a lower number of GPs per capita (Gravelle 1999, Nuscheler 2003), as well as for the substitute relationship between GPs and internists (Newhouse et al. 1982).

Based on our estimations, we employ a decomposition analysis to study the drivers behind the decline in the supply of GPs in Germany over the period 1997–2008. We find that population ageing itself would have supported an increase in supply in both urban and rural regions. In rural regions, this effect is enhanced further by the reduction in population density, leading to a decline in competition. While these impacts are dampened by a concomitant increase in the supply of internists, we find that overall the decline in the number of GPs per capita can only be explained by an increase in their reservation income. This speaks to a (relative) decline in the attractiveness of working in general practice, especially in rural regions.

The remainder of the paper is organized as follows. The next section surveys the related literature, while Section 3 offers a brief introduction into the institutional (Section 3.1) and demographic (Section 3.2) context of the provision of GP services in Germany over the time frame 1995–2009. In Section 4.1 we then develop and solve a theoretical model of regional physician supply, which we employ in Section 4.2 to derive the structural equation and hypotheses. Section 5 presents the empirical analysis, including a description of the data in Section 5.1, our main regression analysis based on different specifications in Sections 5.2 and 5.3, and a decomposition of the change in GP supply based on our empirical findings in Section 5.4. Conclusions are offered in Section 6. Appendix A contains some formal derivations; Appendix B contains additional estimations for robustness, employing absolute numbers of GPs as dependent variable; and Appendix C contains the figures and tables we refer to.

2 Literature

The literature on the geographic distribution of physicians is predominantly empirical and somewhat eclectic in terms of the modelling approach. Newhouse et al. (1982) examine the location patterns of US physicians by regressing the probability of having at least one practice of a given specialty on the size of the town/community. As expected, larger towns are more likely to attract a practitioner within any given specialty. As more specialized physicians require larger catchment areas, they are more prone to settle in larger cities. In turn, this implies that GPs are over-represented in rural areas, where the substitute services of internists are less available. These location patterns have been confirmed by Dionne et al. (1987) for Canadian data, and by Rosenthal et al. (2005) for more recent US data.[3] Their results show that population growth has triggered a diffusion, albeit incomplete, of specialists into more rural areas. While the analytical approach and the results of these studies are well in line with location theory, they are essentially (comparative) static in nature and do not examine the role of population structure or population flows.

A number of studies regress the number of physicians per capita within a certain region on a set of demographic, geographic and economic covariates. The findings with regard to age structure are somewhat mixed. Using cross-sectional data from Switzerland, Kraft and von der Schulenburg (1986) find a positive yet insignificant impact of the population share 55 and above. Using cross-sectional data from Germany, Kopetsch (2007) and Jürges (2007) find a significantly positive impact of the population share 50 and above and 65 and above, respectively.[4] These findings are confirmed by Sundmacher and Ozegowski (2016) in separate estimations for GPs and specialists. Finally, Correia and Veiga (2010) find a very small positive impact of the ageing rate on regional GP supply in Portugal.

In contrast, Hingstman and Boon (1989) identify a significant negative effect of the share of the elderly population on the number of GPs per capita at the regional level within the Netherlands. There, GPs are reimbursed a capitation for each publicly insured patient. This turns elderly people into relatively unprofitable patients as high treatment intensities have to be financed out of a fixed budget per patient. Goddard et al. (2010) estimate the regional physician supply per capita for the UK, using separate cross-sections for 2002 and 2006. They find a negative effect of mean-age, once an IV-estimator is used to address the endogeneity of morbidity. While higher levels of morbidity increase the supply of physicians, older patients may be unattractive prospects nevertheless, as for given levels of morbidity they are potentially more costly to treat.

[3] See also Newhouse (1990) for an 'intermediate' follow-up with US data.

[4] See also Breyer et al. (1986) for an early study on physician supply in Germany. All of these studies estimate physician supply as integral part of an analysis of supplier-induced demand. Breyer et al. (1986) and Kraft and von der Schulenburg (1986) estimate physician supply per capita jointly with measures of expenditure and treatment intensity. Kopetsch (2007) and Jürges (2007) estimate physician supply per capita as an instrument to be used in the estimation of treatment intensity.

The cross-sectional nature of these studies exposes them to the potential problem of unobserved regional heterogeneity. Moreover, it poorly reflects the intertemporal nature of the market for physician services with respect to the evolution of demand.[5] These issues are to some degree addressed in Foster and Gorr (1992), Nocera and Wanzenried (2008) and Gächter et al. (2012), who provide panel data estimations of physician supply for US states, Swiss cantons and Austrian districts, respectively. Nocera and Wanzenried (2008) find that population growth lowers the growth of GP supply but raises the growth of specialist supply, a result that is in line with the argument in Newhouse et al. (1982). However, as the estimation does not directly control for the development of specialist supply when estimating the growth of the number of GPs per capita (and vice versa), it cannot disentangle the pure effect of population growth from the change in competition. Gächter et al. (2012) provide evidence on the interrelationship between private and public GPs and specialists, but none of the studies includes the age structure of the population.

The studies reviewed so far do not develop the estimation framework from the equilibrium condition for physician entry at the regional level. This is only done in Schaumans and Verboven (2008) who extend the Bresnahan and Reiss (1991) model of industry entry to study the complementary relationship between physicians and pharmacies. From the static Nash entry equilibrium they derive structural equations for the probability of certain market configurations, depending on regional market conditions. In their estimation, which is again based on cross-sectional data, they identify a positive impact of the share 65 and above on physician entry.[6]

[5] Correia and Veiga (2010) and Goddard et al. (2010) also provide estimations in differences between two survey years. Correia and Veiga (2010) identify a negative impact of population growth on the growth rate of the per-capita number of GPs, whereas Goddard et al. (2010) find no significant effects for their candidate variables.

[6] Hurley (1991) and Bolduc et al. (1996) provide microeconometric analyses of physicians' choices of their practice location but do not consider the regional population structure as a determinant.

3 Physician Supply and Population Change in Germany 1995–2009

While the relationships between population change, regional structure and physician supply we derive from our theoretical model are entirely general per se, they are importantly shaped by the institutional setting. This is particularly pertinent when it comes to deriving hypotheses on the directions of the relationship between population change and changes in the supply of physicians. We therefore start with a brief introduction into the provision of ambulatory health care (Section 3.1) and population change (Section 3.2) in Germany over the time span 1995–2009 as the context of our empirical analysis.

3.1 Supply of Office-Based Physicians in Germany

In Germany, the group of office-based physicians comprises GPs and specialists (e.g. for internal medicine, gynaecology, ophthalmology, paediatry, orthopaedics, etc.). The large majority of physicians working in ambulatory care are affiliated with the statutory health insurance (SHI), covering about 88 percent of the German population. In 2003, SHI-affiliated physicians (panel doctors) made up for 117,600 out of 132,400 office-based physicians, i.e. for around 89 percent (Busse and Riesberg 2004). The share of GPs amongst all office-based physicians was around 35 percent in 2002.[7] About 75 percent of office-based physicians work single-handed.[8]

The payment for services delivered by ambulatory physicians, including GPs, to SHI-patients is determined in two steps. Initially, sickness funds allocate a negotiated budget to the physician association(s) at the state (Bundesländer) level. In turn, the physician association(s) allocate (notional) budgets to physician practices contingent on the number of patients in the

[7] According to data of the Association of SHI Physicians (Kassenärztliche Bundesvereinigung) the share of SHI-affiliated physicians amongst all family physicians (of which GPs form a substantial part) was 95 percent in 2008. We believe this to be indicative of the share of SHI-affiliated GPs.

[8] For further information on the physician workforce and details of the reimbursement scheme see Busse and Riesberg (2004) and Busse and Blümel (2014).

previous year. Services are then reimbursed on a fee for service basis according to a schedule ("Uniform Value Scale") determined at federal level.[9] One notable feature of the German fee for service system is that once the total volume of reimbursable services has reached a pre-determined budget, any further provision will lead to a proportional reduction in the fee (on all services) such that total payments do not exceed the ceiling.[10] As Busse and Riesberg (2004) report, this has led to almost constant payments per case and per physician despite a sizeable increase in both cases and physicians over the time span 1995–2001.[11] Ambulatory physicians also provide services to the 9 percent of the population with private health insurance. While the remuneration for private patients is typically more generous, the fee structure reflects the one of SHI, yet without budget ceilings. Leaving further details of the payment process aside, two things are worthy of note in regard to reimbursement during the time span 1995–2009 covered by our data:

Observation 1 *(i) Fees for service are typically neither adjusted for age nor for other patient characteristics.*[12] *(ii) While SHI fees may vary across states (Bundesländer), they apply uniformly across all districts (Kreise) within each state, the latter being the geographic unit of observation in our study.*

Patients are free to choose between physicians within each speciality, and although they are encouraged to attend a GP for an initial consultation, there is no strict gatekeeping, implying they can directly access specialists.[13]

[9] Notably, around 21 percent of SHI revenue results from a quarterly "basic fee" per patient that is tantamount to a capitation component.

[10] Initially, budget ceilings were fixed at the level of the regional physician association, which led to a Cournot-type competition of physicians generating an excess incentive to provide services at the expense of a declining price for all, the so so-called "tread-mill effect" (Benstetter and Wambach 2006). From 1997 onwards, budget ceilings were imposed at practice level. While this helped to curb the treadmill effect it also generated an incentive for physicians to shift provision to the privately insured who were not subject to the ceiling (Schmitz 2013).

[11] Busse and Riesberg (2004, Table 32) report an average SHI reimbursement of 171,700 € per GP (236,900 € per internist), which after the deduction of costs leaves a surplus/income of 77,265 € (95,944 €) in 2001.

[12] This excludes the capitation component, which depends on whether or not the patient is retired.

[13] In 2004, a 10 € consultation fee was introduced that was payable at the first consultation with a specific physician within a quarter, consultations upon a referral as well

Indeed, in 1998 around 48 percent of insureds with a large statutory health insurer (AOK) directly consulted a specialist (RKI 2008). Furthermore, for general health care, the age group 0–17 mostly turn to paediatricians even if gradually shifting to GPs with advancing adolescence (Kamtsiuris et al. 2007). From the perspective of GPs, this implies the following:

Observation 2 *(i) With respect to the adult population GPs are in direct competition with specialists for internal medicine (internists), especially if these practise as "family doctors". (ii) The non-adult population (aged 1-17) will contribute relatively little to GPs' income.*

In principle, panel doctors are free to chose their practice location. Since 1993, however, entry regulations apply at the level of physician specialty, including general practice. These regulations are based on the criterion of "excess supply", which up until 2009 was established for a specialty within a district if the respective number of physicians per capita exceeded 110 percent of a benchmark value. The latter corresponded to the 1990 (for GPs and family internists: 1995) number of physicians per capita within this specialty averaged across districts of the same classification.[14] Districts for which an excess supply has been established are generally closed to new entries within the relevant specialty but not to replacements. The opening of a new practice may be granted only if a special need is asserted. While in 2003 a vast majority of districts (typically around 90 percent and more) were classified as exhibiting excess supply for most medical specialties, only 34 percent of districts were classified as exhibiting excess supply for GPs and family internists (Busse and Riesberg 2004). Notably, the numbers of most specialist physicians continued to increase over the 1990s and 2000s despite the almost complete closure of the market that one would expect based on the criterion of excess supply. This suggests a widespread lack of enforcement.

as follow-up consultations with the same physician being free of charge. Evidence on the extent to which this scheme reduced physician visits is mixed (e.g. Schreyögg and Grabka 2010, Farbmacher and Winter 2013).

[14]Notably, this definition was unrelated to patient need, however defined. Neither was it relating to the physician supply within neighbouring regions. See Fülöp et al. (2008, 2010), Klose and Rehbein (2011) and Busse and Blümel (2014) for a more detailed discussion of the German system of entry regulation. The definition of excess supply was altered in the course of a health care reform in 2011, i.e. outside our study period.

We can summarize as follows.

Observation 3 *Although many districts are classified as exhibiting an "excess supply" of physicians, (i) this is much less frequently the case for GPs, and (ii) it does not always result in a binding entry restriction.*

We conclude this section by pointing out a number of changes in the course of health care reforms enacted in 2007 and 2011 (see Busse and Blümel 2014). The first reform provided that from 2009 onwards, the overall budget granted at state level is adjusted for the morbidity of the population and certain regional features, and that volumes of (reimbursable) services are calculated at practice level. The 2011 reform implied changes (i) to the mode of reimbursement, granting greater regional autonomy and introducing various financial and regulatory measures to render rural practice more attractive; and (ii) to the regulation of entry closures, where the share of the population 65+ is now taken into account when determining "need" at district level (Ozegowski and Sundmacher 2012). This reform obviously had no impact on the development of physician supply over our study period.

3.2 Population Change and the Supply of General Practitioners

In the following, we describe some of the salient trends in the supply of GPs at regional level over our period of observation, 1995–2009. We contrast the supply trends against the regional population change over the same time span. The regional INKAR dataset we are employing covers 412 districts (Kreise), corresponding to NUTS III level, as our spatial unit of observation. Districts can be classified into three broad categories: metropolitan districts (type I), municipalized districts (type II), and rural districts (type III).[15]

Over the time span 1995–2009, Germany has witnessed an increase by 27 percent in the total number of office-based physicians (including GPs), while the number of GPs has decreased by about 7 percent. These trends exhibit large variation at district level. Figure 1 provides information on the percentage change in the number of physicians and GPs between 1995 and

[15] We refer the reader to Section 5.1 for a more detailed description of the data.

2009 at district level. While for most districts the number of all physicians exhibits positive growth, the opposite is true for GPs. The latter finding applies in particular to eastern Germany, but also to some rural districts in western Germany. At the same time, some districts in western Germany have experienced an unabated increase in the supply of all physicians and GPs alike.

Figure 1 about here

This begs the question as to what are the causes for these divergent supply trends. One candidate explanation relates to population change, where rural areas have experienced a rapidly shrinking population and a strong increase in the mean age over the observation period. Especially in eastern Germany, the decline in population size has been driven by strong outmigration of young people into metropolitan areas in western Germany that offer better employment prospects.

Population change could potentially affect physician supply through two channels: changes in the age structure of the population and changes in its size. Figure 2 shows that the population share 19minus is negatively correlated both with the population share 60plus (-0.67) and with the share 20–59 (-0.31). Similarly, the population share 60plus is negatively correlated with the population share 20–59 (-0.49). While the share 19minus has decreased by 3.35 percent from 1995 to 2009, the population share 60plus has increased by 5.25 percent during the same period. Hence, the population share 20–59 declined by about 1.9 percent. This is consistent with an ageing population. At the same time, the overall population size declined by about 1 percent.

Figure 2 about here

As Table 2 shows, the three types of districts (metropolitan, municipalized, rural) differ in respect to the nature and speed of population change and in respect to the development of the supply of GPs. While the number of GPs per capita is increasing with the level of rurality, a result which is in line with Newhouse et al. (1982), the number of GPs per square kilometer is decreasing. This is consistent as population density declines even more with rurality. The number of GPs has been in decline over our observation period both per capita and per square kilometer, the decline in numbers being weakest in metropolitan districts.

While the population share 60plus is increasing with rurality by a modest amount, the municipalized districts exhibit the largest share 19minus. Population ageing, as measured by a simultaneous increase in the share 60plus and a decrease in the share 19minus, is manifest across districts of all types. The extent of ageing is most pronounced within rural areas.

Table 2 about here

4 Theoretical Framework

In this section, we develop a theoretical model of entry into a regional physician market in order to determine the number of physicians per capita within entry equilibrium. Based on this, we derive a structural equation, allowing us to regress for each region and at each point in time the per-capita number of physicians on the age structure of the population, on the degree of rurality, on the growth of average list size, on an indicator for entry restrictions, as well as on a set of control variables.

We begin by specifying a physician's income. For this purpose, we focus on some arbitrary region characterized by a time-invariant index of "rurality" r. The resident population of size ℓ_t at time t can be decomposed into three age groups, indexed by $a = 1, 2, 3$. Specifically, we let $a = 1$ correspond to the young age group (19minus), $a = 2$ to the middle age group (20–59), and $a = 3$ to the group of the elderly (60plus). Denoting by ℓ_t^a the size of age group a at time t, we obtain $\ell_t = \sum_a \ell_t^a$ as population size at time t and can then define the age shares $\lambda_t^a := \ell_t^a/\ell_t$. For notational convenience we will write $\lambda_t^1 = \underline{\lambda}_t \in [0, 1]$ and $\lambda_t^3 = \overline{\lambda}_t \in [0, 1]$ so that $\lambda_t^2 = 1 - \underline{\lambda}_t - \overline{\lambda}_t \in [0, 1]$.

We define $y(a, r)$ as a measure of expected income from the provision of ambulatory health care to a resident belonging to age group a within a type-r region and assume[16]

$$y\,(a,r) \geq 0. \tag{A1}$$

Employing the age shares, we obtain

$$\begin{aligned} y_t &= y\left(\overline{\lambda}_t, \underline{\lambda}_t, \zeta_t, r\right) := \sum_a y\,(a,r)\,\lambda_t^a \zeta_t \\ &= \left\{y\,(2,r) + [y\,(3,r) - y\,(2,r)]\,\overline{\lambda}_t + [y\,(1,r) - y\,(2,r)]\,\underline{\lambda}_t\right\}\zeta_t \end{aligned} \tag{1}$$

as the income a physician expects to earn per resident within the region in year t. Here, we use ζ_t to denote a time-variant income shifter, affecting either the demand for the physician's services and/or their profitability. We defer a more concrete definition of ζ_t to further on below [see assumption (A8) in Sub-section 4.2], and merely note at this stage that ζ_t may measure e.g. the availability of substitute medical services, the intensity of competition, regional income and non-age-related compositional measures of the population within the region.[17] Using n_t to denote the number of physicians practising within the region in year t and assuming that physicians serve the local population symmetrically across space and across age groups, we can write a physician's expected period income as $y_t\,(\ell_t/n_t) = y_t\,(n_t/\ell_t)^{-1}$, where ℓ_t/n_t is the (potential) list size of a representative physician in year t.[18]

4.1 Entry Equilibrium

In light of empirical evidence indicating a low geographical mobility of practitioners once they have settled within a particular region (Taylor and Leese 1998, Elliott et al. 2006, Kopetsch and Munz 2007, Correia and Veiga 2010),

[16] This implies either that physicians are able to cross-subisidize the provision of unprofitable services within each age group or, alternatively, that they are able to reduce the intensity of service provision to a level at which they at least break even.

[17] In our estimation ζ_t will be a measure of a number of time-varying control variables for any particular region within our dataset. Thus, ζ_t measures "regional development" rather than "regional type", as the time-invariant index r does.

[18] Noting that the potential list-size is just the inverse of the number of physicians per capita, n_t/ℓ_t, we see that our relationship is consistent with the negative relationship between the supply of GPs per capita and their income that is identified in Dormont and Samson (2008).

the opening of a physician practice can be viewed as a long-term decision.[19] When considering the establishment of a practice, a physician would therefore not only assess the income from the provision of services to the current population, but also the income she expects from the provision of services in the future.

Consider thus a set-up where a representative physician practises for two periods after which she retires or leaves the region/profession for other reasons. In Appendix A1, we set out a more general model in which physicians work for $1 + z$, with $z \geq 1$, periods. We derive our key result (as phrased in Proposition 1) for the general model and show that the two-period case corresponds to the special case $z = 1$.

For the moment, we focus on the physician's life-cycle problem within a single region. Applying a discount factor $\delta < 1$, we can write the present value of the physician's expected life-cycle income at the time of entry, t, as

$$v_t = y_t \frac{\ell_t}{n_t} + \delta E \left(y_{t+1} \frac{\ell_{t+1}}{n_{t+1}} \right)$$

where $y_{\widehat{t}} = y\left(\overline{\lambda}_{\widehat{t}}, \underline{\lambda}_{\widehat{t}}, \zeta_{\widehat{t}}, r\right)$, as given by (1), is the average income per resident in period $\widehat{t} \in \{t, t+1\}$ and where $\ell_{\widehat{t}}/n_{\widehat{t}}$ is the potential list size.[20] The expectations operator for period $t+1$ corresponds to the uncertainty about whether or not the supply of physicians adjusts to its equilibrium value, as described in greater detail below.

Denoting by $u_t = \widehat{u}_t + \delta \widehat{u}_{t+1}$ a non-monetary life-cycle benefit from re-

[19] Unfortunately, there is little data on the average duration of a physician's spell within one specific practice. Elliot et al. (2006) report an average migration rate of 0.012 and an average exit rate of 0.048, adding up to an average turnover rate of 0.06, for GP principals in the UK (year 2003). One issue is that the exit rate includes GPs who enter retirement at the end of their professional life. Combining, thus, the average data in Elliot et al. (2006) with age-specific data on the quitting intentions of English GP principals (year 2001), as reported in Sibbald et al. (2003), one can impute the age structure of actual exit (calculations are available upon request). Based on this, we obtain an expected tenure within a single practice of about 20 years for GP principals in the UK. About 75 percent of the German GPs we are considering work as single-handed principals, and we would presume them to be subject to broadly comparable incentives.

[20] The important role of income in driving a physician's location choice has been established e.g. by Hurley (1991) and by Bolduc et al. (1996) who find (average) income elasticities of 1.05 and 1.11, respectively.

siding in the region, we assume a physician's expected life-cycle utility to be given by[21]

$$w_t = v_t u_t. \tag{2}$$

Furthermore, let $\overline{w}_t$ denote a physician's outside utility at the time t of entering the profession/region and define

$$\overline{v}_t := \overline{w}_t u_t^{-1} \tag{3}$$

as the income required by a physician to take up practice within the particular region. The value of this reservation income increases in the outside utility $\overline{w}_t$ and decreases in the "residential" value of a region u_t.

Within our geographic setting, physicians choose a location from a set of I regions, indexed by $i \in \{1, 2, ..., I\}$. In order to guarantee analytical tractability, we assume (i) that patient mobility across regions is too low as to generate significant cross-regional market overlap, and (ii) that the total number of regions is so large that aggregate physician density $N_t/L_t = \sum_i n_{it}/\sum_i \ell_{it}$ is unresponsive to changes in the number of physicians, n_{it}, within any individual region. The absence of direct spillovers, as by (i), and the absence of sizeable spillovers through changes in the aggregate supply of physicians as a possible determinant of the outside utility $\overline{w}_t$, as by (ii),[22] then implies that the number of physicians within each region i is determined independently from the development in any other region $i' \neq i$.[23] Given the

[21] The multiplicative separability of life-cycle utility w_t into an income stream, v_t, and a stream of non-monetary benefits, u_t, may appear restrictive. It can be shown, however, that a life-cycle utility function of the form $w_t = y_t \frac{\ell_t}{n_t} (\widehat{u}_t)^\alpha + \delta E \left(y_{t+1} \frac{\ell_{t+1}}{n_{t+1}} \right) (\widehat{u}_{t+1})^\alpha$, in which income and non-monetary utility combine period by period leads to an equivalent allocation and ultimately to structural equation (10) under the same set of assumptions (A4)-(A8). A proof is available from the authors on request.

[22] If physicians can be ranked in terms of their outside opportunities, we have $\overline{w}_t = \omega(N_t/L_t, t)$, where $N_t/L_t = \sum_i n_{it}/\sum_i \ell_{it}$ denotes the number of physicians per capita at aggregate level. While assuming that $\omega_{N_t/L_t} \geq 0$ then implies an upward-sloped (inverse) supply function, we maintain that regions are "atomistic" such that isolated changes in n_{it} do not bear on the supply price.

[23] In our empirical analysis we take into account the scope for (unobserved) cross-sectional spillovers and follow Driscoll and Kraay (1998) in calculating error terms that are robust with regard to contemporaneous cross-sectional correlation.

absence of cross-regional effects, we continue to omit the regional index i for the remainder of this section.

Writing the expected life-cycle income in period t as a function of the total number of practising physicians, n_t, in period t and the expected number of physicians n_{t+1} in period $t+1$, respectively, physicians would continue to enter a particular region as long as $v_t(n_t, n_{t+1}) > \overline{v}_t$, i.e. as long as the expected discounted life-cycle income exceeds the reservation income. Assuming the absence of entry restrictions, a period t entry equilibrium is then given by the number of physicians n_t^* that satisfies $v_t(n_t^*, n_{t+1}) = \overline{v}_t$.[24] However, it is not clear a priori whether physician turnover by entry and/or exit leads to an adjustment towards the equilibrium number of physicians $n_{\hat{t}}^*$ in all periods. Disequilibrium situations may arise from constraints on both entry and exit. Under a binding entry restriction the number of practising physicians falls short of the equilibrium number, such that $n_{\hat{t}} = \overline{n} < n_{\hat{t}}^*$ and $v_{\hat{t}}(\overline{n}, n_{\hat{t}+1}) > v_{\hat{t}}(n_{\hat{t}}^*, n_{\hat{t}+1}) = \overline{v}_{\hat{t}}$. Conversely, the presence of sunk costs may lead to a situation where physicians do not exit the market before their retirement although the number of practising physicians exceeds the equilibrium number, implying that $n_{\hat{t}} = \underline{n} > n_{\hat{t}}^*$ and $v_{\hat{t}}(\underline{n}, n_{\hat{t}+1}) < v_{\hat{t}}(n_{\hat{t}}^*, n_{\hat{t}+1}) = \overline{v}_{\hat{t}}$.

We will, thus, derive the equilibrium supply n_t^* of physicians in some period t, depending on whether or not physician supply adjusts to changes in the regional environment in the subsequent period $t+1$. Specifically, we assume that physicians face uncertainty at time t as to whether or not an entry equilibrium in period $t+1$ will be realized. Let $s \in [0,1]$ denote the probability a physician assigns in period t to facing a disequilibrium in period $t+1$ and let

$$\psi_{t+1} \in \begin{cases} \left[1, n_{t+1}^*/n_t^*\right] & \textit{for } n_{t+1}^*/n_t^* > 1 \\ \left[n_{t+1}^*/n_t^*, 1\right] & \textit{for } n_{t+1}^*/n_t^* < 1 \end{cases} \tag{4}$$

denote the extent of adjustment towards the equilibrium value n_{t+1}^* in a disequilibrium situation. The expected income per resident in period $t+1$

[24]Stability is readily verified, as $\partial v_t / \partial n_t = -y_t\left(\ell_t / n_t^2\right) < 0$. For simplicity, we ignore the integer issue.

can then be written as

$$E\left(y_{t+1}\frac{\ell_{t+1}}{n_{t+1}}\right) = sy_{t+1}\frac{\ell_{t+1}}{\psi_{t+1}n_t^*} + (1-s)\,y_{t+1}\frac{\ell_{t+1}}{n_{t+1}^*}$$

with $\psi_{t+1}n_t^*$ denoting the number of physicians in period $t+1$ in a disequilibrium situation.

Assume that the discount factor δ is sufficiently small such that

$$\delta < \frac{1}{(1-s)\,\overline{v}_{t+1}/\overline{v}_t} \tag{A2}$$

is satisfied.[25] In Appendix A1, we then prove the following:

Proposition 1 *(i) Within an entry equilibrium at time t the number of physicians per capita is (approximately) given by*

$$\frac{n_t^*}{\ell_t} \approx \frac{y_t + s\delta y_{t+1}\frac{\ell_{t+1}}{\psi_{t+1}\ell_t}}{\overline{v}_t - (1-s)\,\delta\overline{v}_{t+1}}. \tag{5}$$

(ii) Expected restrictions to entry (exit) in period $t+1$ imply an over-supply (under-supply) of physicians in period t relative to the supply $\frac{n_t^*}{\ell_t}\,|_{s=0} = \frac{y_t}{\overline{v}_t - \delta\overline{v}_{t+1}}$ that is supported by the current economic and demographic structure, i.e. $n_{t+1} = \psi_{t+1}n_t^* < (>)\, n_{t+1}^* \Leftrightarrow \frac{n_t^*}{\ell_t}\,|_{s>0} > (<)\, \frac{n_t^*}{\ell_t}\,|_{s=0}$.

According to part (i) of the Proposition, the equilibrium number of physicians per capita at time t always increases with the income y_t per resident in this period and decreases with the reservation income $\overline{v}_t$ at the point of entry. The extent to which the (prospective) income y_{t+1} per resident in period $t+1$ determines the number of physicians in period t depends on the expectation about whether or not the number of physicians in period $t+1$ will adjust to its equilibrium value. In case it does with certainty such that $s = 0$, surprisingly perhaps, the prospective income has no bearing on the current number of physicians per capita. This is because the income stream in period $t+1$ is fully offset by adjustments in the equilibrium number of physicians, n_{t+1}^*.

[25] From our estimations we obtain $s \geq 0.481$ (see Section 5.4) and $\overline{v}_{t+1}/\overline{v}_t < 1.121$ (see Table 5, row (8)). Hence, $(1-s)\,\overline{v}_{t+1}/\overline{v}_t < 1$. Given that we may reasonably assume that $\delta \leq 1$, the inequality in (A2) is satisfied.

For a physician pondering entry at period t this implies that the discounted value of the income stream during the second period of her working life equals approximately the discounted value of the reservation income $\delta\overline{v}_{t+1}$ at $t+1$. Thus, while the number of physicians per capita n_t^*/ℓ_t increases with this value, it is no longer responsive to the income stream in period $t+1$. Any expected increase in demand, for instance, is offset by a corresponding increase in the number of rivals.

If, in contrast, physician supply adjusts only partially such that $s=1$, then the current number of physicians n_t^* increases with the discounted level of prospective income y_{t+1} and with the growth in the list size

$$\frac{\ell_{t+1}}{\psi_{t+1}\ell_t} = \frac{\ell_{t+1}/n_{t+1}}{\ell_t/n_t^*} = 1 + g_{t+1}^{\ell/n} \tag{6}$$

under incomplete adjustment. Since $1+g_{t+1}^{\ell/n} = \frac{1+g_{t+1}^{\ell}}{1+g_{t+1}^{n}}$, with g_{t+1}^{ℓ} and g_{t+1}^{n} denoting the growth rates of the population ℓ and the number of practitioners n, respectively, it is easy to see then that population change has an impact on the current number of physicians per capita if and only if there is imperfect adjustment to the entry equilibrium.

According to part (ii) of the Proposition, future entry restrictions are anticipated in the current supply of physicians. An entry restriction in period $t+1$ implies excessive list size growth relative to a situation of perfect equilibrium adjustment.[26] The expectation of this then triggers additional entries in period t beyond the number of physicians per capita that would be supported by the current economic and demographic structure. The converse is true if exit restrictions lead to an excessive shrinking of the list size for a declining population.

4.2 Structural Equation and Hypotheses

While our theoretical results for the two-period model are entirely general in qualitative terms, they imply a period length of around 20 years. This is, of course, an abstraction from reality, where turnover in the physician market takes place at a higher rate and is typically measured in yearly intervals.

[26] This is readily checked when setting $\psi_{t+1} < n_{t+1}^*/n_t^*$ in (6).

Accordingly, our data are based on yearly observations, stretching over a period of 15 years. Hence, our structural equation should be based on a period length of 1 year. When deriving Proposition 1 in Appendix A1, we consider a general set-up, where a representative physician practises for $1+z$, with $z \geq 1$, periods after which she retires or leaves the region/profession for other reasons.

Defining $g^{y}_{it+1} := y_{it+1}/y_{it} - 1$ and $g^{\overline{v}}_{it+1} := \overline{v}_{it+1}/\overline{v}_{it} - 1$ as the growth rate of the (expected) income earned per resident and of the reservation income, respectively, and assuming[27]

$$\delta \ll \min\left\{s^{-1}, \left[(1-s)\left(1+g^{\overline{v}}_{it+1}\right)\right]^{-1}\right\}, \tag{A2'}$$

$$z \gg 1, \tag{A3}$$

we can show that the equilibrium number of physicians per capita in region $i \in \{1, 2, ..., I\}$ at time t is approximated by

$$\frac{n^*_{it}}{\ell_{it}} \approx \frac{y_{it}\left[1 + s\delta\left(1+g^{y}_{it+1}\right)\left(1+g^{\ell/n}_{it+1}\right)\right]}{\overline{v}_{it}\left[1-(1-s)\,\delta\left(1+g^{\overline{v}}_{it+1}\right)\right]}; \quad s \in [0,1], \tag{7}$$

which is equivalent to equation (5) in Proposition 1. Intuitively, the two-period model constitutes a good approximation of the general $1+z$ period model whenever there is (i) a sufficient extent of discounting, implying that effects of a changing practice income associated with the expectation of entry constraints in a period $\widehat{t} > t+1$ are valued at a factor $(s\delta)^{\widehat{t}-t}$ small enough to be of second order; and (ii) a sufficient duration of the remaining working life z, implying that changes to physician income after the retirement of the current entrants can be disregarded.

In preparation of our empirical analysis, we can then transform (7) into a structural equation and formulate hypotheses about the determinants of regional physician supply. To this end, define

[27] By $x \ll (\gg)\, y$ we denote that x is much smaller (larger) than y.

$$\overline{\beta}(r_i) \ := \frac{[y(3,r_i) - y(2,r_i)]}{y(2,r_i)}, \tag{8}$$

$$\underline{\beta}(r_i) \ := \frac{[y(1,r_i) - y(2,r_i)]}{y(2,r_i)}, \tag{9}$$

and assume

$$g^{y}_{it+1} \rightarrow 0, \tag{A4}$$

$$g^{\overline{v}}_{it+1} \approx g^{\overline{v}}, \tag{A5}$$

$$u_{it} \approx u_i \tag{A6}$$

$$\overline{\beta}(r_i)\,\overline{\lambda}_{it} + \underline{\beta}(r_i)\,\underline{\lambda}_{it} \ll 1, \tag{A7}$$

$$\zeta_{it} = \prod_h \overline{\zeta}_i \overline{\zeta}_t \left(control_{hit}\right)^{\gamma_h}, \quad \overline{\zeta}_i, \overline{\zeta}_t > 0. \tag{A8}$$

Assumption (A4) implies that the expected income a physician earns per resident does not vary too much between periods; (A5) implies that the growth rate of the outside utility, reflecting e.g. exogenous supply trends, is approximately constant over time and across regions; (A6) implies that the benefit from regional amenities is approximately constant over the time span under consideration; (A7) implies that the weighted sum of the age shares is sufficiently smaller than one; (A8) implies that the income shifters (= region-time specific control variables) follow a Cobb-Douglas function.

Based on these assumptions, we derive in Appendix A2 the following structural equation[28]

$$\ln\left(\frac{n^*_{it}}{\ell_{it}}\right) = \beta_0 + \beta_i + \beta_t + \overline{\beta}(r_i)\,\overline{\lambda}_{it} + \underline{\beta}(r_i)\,\underline{\lambda}_{it} + \delta'\left(1 + g^{\ell/n}_{it+1}\right) + \sum_h \gamma_h \ln\left(control_{hit}\right) \tag{10}$$

[28]In the course of our empirical analysis, we verify that the key assumptions (A2'), (A4) and (A7) are compatible with our estimation results.

with

$$\begin{aligned}
\beta_0 &: = -\ln\left[1-(1-s)\,\delta\left(1+g^{\overline{v}}\right)\right], \\
\beta_i &: = \ln y\,(2,r_i) + \ln\overline{\zeta}_i + \ln u_i, \\
\beta_t &: = \ln\overline{\zeta}_t - \ln\overline{w}_t, \\
\delta' &: = s\delta.
\end{aligned}$$

Hence, the logarithm of the number of physicians per capita depends on (i) a region-fixed effect β_i; (ii) a time trend β_t; (iii) the short-run effects of the age-shares according to the coefficients $\overline{\beta}\,(r_i)$ and $\underline{\beta}\,(r_i)$; (iv) the growth factor in period $t+1$ of the average list size according to the coefficient δ'; and (v) a number of logarithmic control variables. The region-fixed effect embraces in particular the impact of (region-specific) profitability $y\,(2,r_i)$ of a member of the population aged 20–59 living in a region with a degree of rurality r_i, a region-specific impact on demand $\overline{\zeta}_i$, and the benefit u_i from residing within a region i.[29] The time trend relates in particular to physician income, reflecting changes in the level of demand or adjustments in reimbursement, as well as to the trend in the outside utility of becoming a physician, including changes in aggregate supply.[30]

The coefficients on the age shares $\overline{\beta}\,(r_i)$ and $\underline{\beta}\,(r_i)$ measure the percentage change in profitability if a member of the population 60plus or 19minus, respectively, is treated rather than a member of the age group 20–59. Obviously, the sign of these coefficients is positive if and only if the treatment of old or young members of the population is more profitable than the treatment

[29] Alternatively to (A6), we could assume that the benefit from regional amenities $u_{it} = \prod_k \overline{u}_i\,(control_{kit})^{\mu_k}$ depends on a set of time-variant control variables. Following through the derivation in Appendix A2, we would then obtain $\beta_i := \ln y\,(2,r_i) + \ln\overline{\zeta}_i + \ln\overline{u}_i$ and an additional term $\sum_k \mu_k \ln(control_{kit})$ in (10). Some of the control variables we are using (e.g. GDP per capita, share of foreigners, share of school leavers with higher education qualifications, tourist accommodation per 100,000 population, share of in-commuters in total employment) are then reflecting both time-variant regional amenities and demand shifters. We would maintain, however, that our main variables of interest, the age shares and list size growth, are predominantly working through the income channel.

[30] Note that from an econometric perspective, the impact in the aggregate supply of physicians per capita (across all regions), N_t/L_t, is effectively indistinguishable from the "pure" time trend β_t. Indeed, we find in our data that both the N_t/L_t and the lagged value N_{t-1}/L_{t-1} are perfectly correlated with the time trend in our data. We therefore do not include measures of aggregate supply as distinct controls.

of middle-aged persons.

We have argued earlier that this is likely to depend on whether a region is urban or rural as well as on the extent to which differences according to age and location in the profitability of treating individuals are balanced out (or magnified) through the payment system. To illustrate the relationship, we express the expected income from providing services to a resident of a type-r_i region who belongs to age group a as

$$y(a, r_i) = \pi(a, r_i)\,\theta(a, r_i),$$

where $\theta(a, r_i)$ and $\pi(a, r_i)$ measure the expected demand for services and the average mark-up on the provision of services to patients aged a within a type-r_i region, respectively. The latter can be written as, $\pi(a, r_i) := \sum_k x(a, r_i, k)\,\pi(a, r_i, k)$, with $\sum_k x(a, r_i, k) = 1$, and amounts to a weighted average of the mark-up $\pi(a, r_i, k) = p(a, r_i, k) - c(a, r_i, k)$ on a set of specific services, with $p(a, r_i, k)$ and $c(a, r_i, k)$ denoting the unit fee and unit cost for the provision of service k to a member of age group a in a region of type r_i. As is readily seen age and regional type then impact on the average mark-up through variations in the service mix, i.e. the shares $x(a, r_i, k)$, as well as through the structure of the fees and costs for the various services.

Writing the coefficient of the age-share 60plus as

$$\overline{\beta}(r_i) = \frac{y(3, r_i)}{y(2, r_i)} - 1 = \frac{\pi(3, r_i)}{\pi(2, r_i)}\frac{\theta(3, r_i)}{\theta(2, r_i)} - 1,$$

we see that the provision of health care to the population 60plus is relatively profitable if they demand more services (to sufficient extent), such that $\frac{\theta(3,r_i)}{\theta(2,r_i)} > 1$, or if the provision of services to them commands a (sufficiently) higher average markup, such that $\frac{\pi(3,r_i)}{\pi(2,r_i)} > 1$. The average mark-up for the provision of services to the population above 60 tends to be larger if they consume larger shares of profitable services and/or if they command a larger mark-up on a given set of services. The latter depends on the extent to which fees are adjusted to reflect differences in treatment costs across age groups and regional context. A similar argument applies for the relative profitability of the age group 19minus, as given by $\underline{\beta}(r_i)$.

Given the lack of detailed data on both the (a, r_i)-specific demand for services and the (a, r_i)-specific average mark-up, we directly estimate the relative profitability of the different age groups depending on regional type. To this end, we let $\overline{\beta}(r_i) = \overline{\beta} + \overline{\beta}_{r_i}$ and $\underline{\beta}(r_i) = \underline{\beta} + \underline{\beta}_{r_i}$, respectively, where $\overline{\beta}_{r_i}$ and $\underline{\beta}_{r_i}$ are dummy variables depending on a region's type r_i. In line with our data, we consider three regional types $r_i \in \{I, II, III\}$, the roman numbers reflecting increasing levels of rurality. Using the most urban type of region as reference, we let $\overline{\beta}_I = \underline{\beta}_I = 0$, implying that $\overline{\beta}$ and $\underline{\beta}$ measure the impact of the age shares within metropolitan areas.

While in principle our model allows for a general set of predictions, the previous argument has shown that the relative profitability of different population groups depends on both regional context and the reimbursement scheme (as well as on other features of the health care system). Where relevant, we therefore formulate our hypothesis against the specific context of the provision of GPs services in Germany, 1995–2009.

H1 $\underline{\beta} < 0$ for GPs who are the focus of our empirical analysis.

The first hypothesis is easily related. Recall from part (ii) of Observation 2 in Section 3.1 that in Germany, the age group 19minus is typically served by paediatricians or other specialists, leaving little demand for GPs, so that $y(1, r) \ll y(2, r)$ in (9), which implies the negative sign. As the population aged 19minus is unprofitable relative to the middle-aged population, a greater share $\underline{\lambda}_t$ will ceteris paribus lower a GP's expected income.

H2 (a) $\overline{\beta} > 0$. (b) $\overline{\beta}_{III} < \overline{\beta}_{II} < 0$.

As we have argued earlier, the impact of the population share 60plus on the supply of physicians depends on the regional make-up as well as on the extent to which the reimbursement system adjusts for age- and location-related differences in the demand and mix for treatments and their costs. Register data of a large statutory health insurer (AOK), covering about a third of the German population, shows for the year 2008 that both utilization of ambulatory care and per capita spending on ambulatory care tend to increase over the age-range 19–80 (Gerste 2012). Although receding for the highest ages, both utilization and spending are unambiguously higher for patients drawn from the age group 60plus as opposed to the age group 20–59.[31] Part (i) of Observation 1, according to which a large part of a GP's reimbursement is not adjusted for patient age nor for practice location, then suggests an ambivalent effect: While we would expect the population 60plus to exhibit a larger demand as opposed to the middle age-group, whether or not the mark-up on services towards this age-groups is higher or lower crucially depends on the mix of services and on their cost. Unfortunately, we have no data on how these vary with age and with the degree of location.

A reading of the literature suggests there may be a regional gradient to the relative profitability of the population share 60plus: Specifically, we conjecture that while a greater share of the elderly population raises physician income and, thus, the per-capita number of physicians, in metropolitan regions ($r_i = I$),[32] this effect is diminished, and possibly reversed, in more rural regions ($r_i \in \{II, III\}$). While we thus assume that $y\,(3, I) > y\,(2, I)$, more elderly populations may be less profitable in rural settings for at least two reasons. First, while longer travelling times generally restrict the access to services and, thus, effective demand within rural regions in general (e.g. Dusheiko et al. 2002, Arcury et al. 2005, Iversen and Kopperud 2005, Thode et al. 2005), this effect is typically more pronounced for the elderly (Chaix et al. 2005, Ono et al. 2014).[33] Second, the provision of services to an older patient may be relatively less profitable within a rural area. Arcury et al. (2005) show that the composition of physician services consumed by an individual switches from preventive and chronic care to acute care as travel distance/time increases. In particular, home visits and out-of-hours care are prone to play a much larger role in rural environments.[34] On the one hand,

[31] These observational findings are confirmed by regression analysis by Jürges (2009), based on individual data, and Kopetsch and Schmitz (2014), based on district-level data. According to these studies the utilization of office-based physician services tends to increase with age from middle ages onwards as well as with various indicators of bad health or medical need.

[32] Since $\overline{\beta}_I = 0$, we then must have $\overline{\beta} > 0$.

[33] This does not rule out that elderly people have a higher demand for GP services in all areas. The focus here is on the relative effect of rurality.

[34] In 1994, the share of home visits in all face-to-face consultations made up 9 percent

home visits are much more frequent both amongst the elderly as opposed to the middle-aged population and within rural regions (e.g. Aylin et al., 1996; Straand and Sandvik, 1998; Boerma and Groenewegen, 2001; Giuffrida and Gravelle, 2001). On the other hand, home visits are typically less profitable than office-based services (Boerma and Groenewegen, 2001). We would then expect an increasing degree of rurality to lower the profits earned on treating an elderly as opposed to a middle-aged population because the composition of services is shifted towards the relatively unprofitable home visits, and in particular so among the elderly. Furthermore, longer travelling times within rural regions tend to lower the profitability of home visits and, thus, the profitability of the service that has a high incidence among elderly patients. The lack of regional adjustments in the reimbursement of office-based physicians [see Observation 1, part (ii)] suggests that these arguments apply very well to Germany.[35]

H3 $\delta' > 0$ if and only if the adjustment to equilibrium is incomplete across a sufficient number of regions.

A significant positive coefficient on list size growth implies that there must be restrictions to equilibrium adjustment within a substantial number of local physician markets. Here, an estimate $\delta' = s\delta > 0$ is reflecting both the expected or average degree of adjustment $s \in [0, 1]$ across all regions and the discount factor. From Observation 3 we would, indeed, expect an interior level of $0 < s < 1$ for Germany due to the presence of partial entry controls.

We can gain further leeway by using information on whether or not specific regions were classified as exhibiting "excess supply" in period t and should therefore have been subject to (unobservable) entry controls in period $t + 1$. More specifically, we estimate a version of the model, with the

in Germany (Marshall, 1996).

[35] In 2002 a (scheduled) home visit was reimbursed in Germany at the same point value (300) as an (intensive) diagnostic and therapeutic consultation $<$ 30 min. An emergency home visit was reimbursed at 600 points and, thus, at the same level as a consultation $>$ 30 min (see Busse and Riesberg; 2004: table 30). Given the travel and time costs involved with a home visit, this suggests that an office-based provision is more profitable. Furthermore, as the reimbursement value of home visits does not increase with travelling time/distance, we would expect the profitability of this service to decline with the level of rurality.

coefficient on list-size growth, $1+g^{\ell/n}_{it+1}$, given by $\delta_{it} = \delta' + \delta'' s_{it}$, where $s_{it} = 1$ if there was excess supply in region i at time t and $s_{it} = 0$ otherwise. From this, we obtain the additional hypothesis:

H3' $\delta'' > 0$.

If an entry ceiling in period $t+1$ imposes a restriction on the adjustment of physician supply, this should imply that list size growth in period $t+1$ is anticipated in a greater supply of physicians in period t. Note, however, that we cannot rule out $\delta_{it} = \delta' > 0$ for regions which are not subject to regulatory entry restrictions ($s_{it} = 0$), as equilibrium adjustment may be limited for non-regulatory reasons. Nevertheless, the impact of list size growth should be stronger for regions which are (potentially) subject to regulatory entry ceilings.

The coefficients γ_h on the control variables measure the effects of the profit shifters. Here, population density (residents per square kilometer), and the numbers of internists and paediatricians per capita, respectively, are of particular interest.

H4 $\gamma_{popdens} < 0$.

According to theoretical models of spatial competition (e.g. Salop 1979, Gravelle 1999, Nuscheler 2003) the intensity of competition tends to increase with population density.[36] For the particular case of Germany and over the period of our observation, this may well have come in the form of "treadmill competition" with the provision of additional services eroding the revenue of all physicians due to budget-balancing reductions in the fee across the board (Benstetter and Wambach 2006). With competition thus leading to

[36] Gravelle (1999) considers a set-up where ℓ (potential) patients are uniformly distributed around a circle of circumference K, implying a population density ℓ/K at each point of the circle. Physicians who locate (symmetrically) around the circle then compete for the marginal set of patients, as given by ℓ/K. The higher the density the more patients can thus be gained by a marginal reduction in the fee charged or, equivalently, by a marginal increase in quality. Put in different terms, the demand elasticity tends to be larger in more densely populated areas (Salop 1979). In equilibrium, this leads to lower prices and/or higher quality and to lower operating profit.

an erosion of income, more densely populated areas sustain fewer competitors for a given population size.

H5 $\gamma_{intern\_pc} < 0$.

It is well-known from Newhouse et al. (1982) that the supply of GPs is negatively correlated with the supply of internists, as these two groups are providing substitute services. This implies hypothesis H5.[37] Recall from part (i) of Observation 2 that this is all the more true as in Germany a significant share of internists have joined GPs in taking on a function of "family practitioner".

H6 $\gamma_{paed\_pc} > 0$.

The role of paediatricians is less clear-cut. A priori, one might expect a negative correlation between the supply of paediatricians and the supply of GPs, as they are also providing substitute services. However, in contrast to internists, the competition between GPs and paediatricians is concentrated on the age group 19minus alone, implying that this effect would be comparatively weak from a population perspective. Note that this is in line with our assertion that GPs are only making comparatively low profits based on the treatment of members of the population aged 19minus, $y(1, r) \ll y(2, r)$. In light of this, a good provision of paediatric services may rather be an indicator of regional attractiveness, especially for young GPs who are in the process of choosing a practice location within a family-friendly environment. And this is what we hypothesize in H6.

We conclude by noting that the effects relating to internists and paediatricians should not be interpreted as causal, as the supply of all three groups of physicians is typically determined simultaneously.

[37] From Newhouse et al. (1982) it is also known that the supply of internists tends to increase with population density. Note, however, that as we are controlling for both population density and for the number of internists per capita, the coefficients $\gamma_{popdens} < 0$ and $\gamma_{intern\_pc} < 0$ report the appropriate partial effects (holding the respective other factor constant).

5 Empirical Analysis

5.1 Data

We build our empirical analysis on regional panel data provided by the Federal Office for Civil Engineering and Regional Development (BBR) as part of the INKAR data set. This data covers 412 districts (Kreise) in Germany (corresponding to NUTS III level) over the time span 1995–2009. Due to the unbalanced structure of the data matrix and the calculation of further variables (growth rates, leads) we consider in our regressions 409 districts and the period 1997–2008.

We use the number of SHI-affiliated GPs per 1000 residents at district level as our dependent variable.[38] We focus on GPs as they constitute a large and homogeneous group of physicians who practise mostly independently from particular features of a district, such as the presence or absence of a hospital, and are not subject to speciality-specific payment arrangements and regulations. GPs typically provide services to a very local population, implying that cross-district spillovers are of little relevance.[39] Finally, they are less prone to be subject to regulatory entry ceilings [see Observation 3, part (i) in Section 3.1], implying that we should observe sufficient variation in the endogenous variable.

Demographic change may bear on the supply of physicians by way of changes to both the age structure and the size of the population (Hypotheses H1–H3). We approximate the age structure by including both the population share of people 60 years and older (60plus) and the population share of people under 20 years (19minus). In order to test Hypothesis H2 that the impact of age structure on physician supply may depend on the district's character, we interact both age groups (60plus and 19minus) with an index of rurality, as provided by the BBR. This index is subaggregated into metropolitan districts

[38] As we have outlined in footnote 7, around 95% of GPs hold an SHI-affiliation. With the INKAR dataset containing the universe of SHI-physicians, we are confident our data covers around 95% of all GPs.

[39] According to a simulation analysis based on utilization data for German ambulatory health care, 71.5% of the German population visit the GP closest to their residential location (Fülöp et al. 2011). Similar findings apply for the Nehtherlands and Belgium (Schaumans and Verboven 2008).

(type I), municipalized districts (type II), and rural districts (type III). The advantage of this index is that it aggregates information on population and possible spillover effects from supraregional cities.[40]

Unfortunately, we are lacking data about the reimbursement schedule as an important intermediary between population structure and regional character as drivers of physician supply. Recalling from part (ii) of Observation 1 that fees are determined at state level without further adjustments to district features, we include state-year effects in our estimation, which control for variation in the fees across states and over time.[41] Another important determinant of income is the share of population with private health insurance. Lacking data on the distribution of insurance status among the population at district level, we would assume this to be picked up by the district-fixed effects and possibly by the variable on income per capita at district level, which we conjecture to be positively related to the share of the privately insured.[42]

According to Hypothesis H3, the growth factor of the average list size within a district constitutes a crucial measure for the extent to which the future development of practice income has a bearing on the current supply of physicians. We thus include in our estimation the list size growth factor advanced by one period. According to the extended Hypothesis H3', the interaction between list size growth and a measure of regulatory entry ceilings should have an additional positive impact. Unfortunately, district-level data are available neither on the presence of entry closures (by specialty) nor on exceptional permissions. We can, however, construct from the data a dummy

[40] Type I districts are defined by more than 300,000 inhabitants or a population density of at least 300 inhabitants/km^2. Type II districts are defined by a population density of at least 150 inhabitants/km^2 or at least 100,000 inhabitants and a population density of at least 100 inhabitants/km^2. Type III districts are defined by a population density of at least 150 inhabitants/km^2 with less than 100,000 inhabitants or at least 100,000 inhabitants and a population density of less than 100 inhabitants/km^2.

[41] As we have detailed in Section 3.1, the 2007 and 2011 health care reforms should not have a bearing on our estimation except for the year 2009, when the morbidity adjustment of budgets at state level was first introduced. Any impact this may have had would be captured by the state-year effect.

[42] Recent cross-sectional studies by Sundmacher and Ozegowski (2016) and Vogt (2016) employ a constructed measure of the share of the privately insured in the year 2010. We cannot use this particular variable in our panel analysis, as it is perfectly correlated with the district-fixed effects.

variable for each district at each point in time, which is set to one when the criterion of "excess supply" is satisfied and set to zero otherwise. As discussed previously, "excess supply" is defined by the number of physicians per capita in a district exceeding 110 percent of a benchmark value (by specialty and district type). We take the benchmark values for family doctors (Hausärzte) as reported in Klose and Rehbein (2011, Table 2).[43] One complication is that the category "family doctors" for which an excess supply of doctors is determined includes GPs and family-oriented internists as a sub-category of all internists. Unfortunately, for our observational period 1995–2009, we neither have data on the number of family doctors nor on the number of family-oriented internists. For this reason, we are unable to construct in a precise way the indicator for "excess supply" and resort to considering an upper and a lower bound: For the upper bound, we contrast the number of "GPs plus all internists" per capita against the benchmark for family doctors, implying that we are establishing "excess supply" in more cases than there were in reality. For the lower bound, we contrast the number of GPs per capita against the benchmark for family doctors, implying that we are assigning "excess supply" to fewer cases than there were in reality. In section 5.3, we present estimation results for both proxies and show that they do not differ much.

In order to test Hypothesis H4, we directly control for population density. Finally, in order to account for the interrelation between GPs and internists (Hypothesis H5) and paediatricians (Hypothesis H6), we employ the per-capita numbers of internists and paediatricians at district level as control variables. The remaining set of control variables comprises: GDP per capita in 1,000 €, share of foreigners within the population, share of school leavers with higher education entrance qualification, tourist accommodation per 100,000 population, and the share of in-commuters within total employment, all variables taken at district level. These variables allow us to control for other time-varying impacts on physician density, as may be transmitted both through income and through the non-pecuniary benefits from taking up

[43]Note that the values are reported as "residents per physician" and, thus, need to be inverted.

practice within a district.[44] Table 1 provides a summary of the variables we employ.

Table 1 about here

5.2 Demographic and Geographic Determinants of GP Supply: Estimation and Results

We now proceed to establish for a representative district the empirical relationship between the number of GPs per capita on the one hand, and the demographic and geographic make-up of the district on the other. In this section, we present findings for a model in which we do not yet account for the role of the "excess supply" criterion, deferring this analysis to the following Subsection 5.3. Following the structural equation (10), the general specification of our panel data model for GP supply is

$$\begin{aligned}\ln\left(\frac{n^*_{ijt}}{\ell_{ijt}}\right) &= \beta_0 + \overline{\beta}\,\overline{\lambda}_{ijt} + \overline{\beta}_{II} ruralII_{ij}\overline{\lambda}_{ijt} + \overline{\beta}_{III} ruralIII_{ij}\overline{\lambda}_{ijt} \\ &+ \underline{\beta}\underline{\lambda}_{ijt} + \underline{\beta}_{II} ruralII_{ij}\underline{\lambda}_{ijt} + \underline{\beta}_{III} ruralIII_{ij}\underline{\lambda}_{ijt} + \delta' listgro_{ijt+1} \\ &+ \gamma_1 \ln popdens_{ijt} + \gamma_2 \ln intern\_pc_{ijt} + \gamma_3 \ln paed\_pc_{ijt} \\ &+ \sum_{h=4}^{H} \gamma_h \ln control_{hijt} + \sum_j \beta_{jt} B_{jt} + \beta_i + \epsilon_{ijt} \qquad (11)\end{aligned}$$

with $i = 1, 2, ..., I$ districts, $j = 1, 2, ..., J$ states and $t = 1, 2, ..., T$ periods. Here, $\ln\left(n^*_{ijt}/\ell_{ijt}\right)$ denotes the logarithm of the number of GPs per capita, while $\overline{\lambda}_{ijt}$ and $\underline{\lambda}_{ijt}$ denote the population share 60plus and share 19minus, respectively, the reference group being the share 20–59. The terms $ruralII_{ij} \times \overline{\lambda}_{ijt}$ and $ruralIII_{ij} \times \overline{\lambda}_{ijt}$ measure the interaction of the population share 60plus with the rurality levels II (municipalized districts) and III (rural districts), a metropolitan district being the reference type. Analogous

[44] Unfortunately, the INKAR dataset does not include data on the gender composition of the population nor on morbidity indicators. Therefore, and in line with other (even just cross-sectional) studies using this data (e.g. Scholz et al. 2015, Sundmacher and Ozegowski 2016, Vogt 2016), we are unable to control for these effects.

notation applies to the interaction of rurality with the population share 19minus. The variable $listgro_{ijt+1}$ denotes the growth factor of the average list size advanced by one year, $popdens_{ijt}$ is the population density, $intern\_pc_{ijt}$ and $paed\_pc_{ijt}$ are the numbers of internists and paediatricians per capita, respectively, and $control_{hijt}$ is a vector of $h = 4, 5, ..., H$ additional control variables. Furthermore, d_{jt} are state-year effects, β_i are district-fixed effects, and ϵ_{ijt} is an error term. The term β_{jt} captures general time trends in the supply of GPs and variation in reimbursement that is determined at the state level (with B_{jt} denoting the corresponding dummies). Recall that German states (Bundesländer) comprise a whole number of districts (Kreise).

We calculate standard errors that are robust to heteroskedasticity and to contemporaneous cross-sectional correlations in the error terms, following Driscoll and Kraay (1998). According to these authors spatial correlations among cross-sections may arise for a number of reasons, ranging from observable common economic shocks to unobserved contagion or neighbourhood effects.

Table 3 provides different specifications of Equation (11).[45] Our preferred specifications are (4) and (5).[46] In regressions (5), (7), (9), and (11) we consider the interaction of the population shares with the dummies for type II and III districts, respectively. In regressions (6) and (7) we omit the additional controls, while in regressions (8) and (9) we exclude the numbers of internists and paediatricians per capita. In regressions (10) and (11) we exclude both sets of variables in order to assess their role in shaping the relationship between the demographic and geographic variables and GP supply.

Table 3 about here

For all regressions we find that a higher share of the elderly population affects the number of GPs per capita positively. The additional interactions with rurality levels II and III show, however, that the positive effect of the

[45] In Table 3 we also provide the results of the Maddala-Wu panel data unit root test. Although the test clearly rejects the null hypothesis, we should not overrate this because of the relatively short time frame 1997–2008. Complete results are available upon request.

[46] See Table 7 in Appendix C for the complete set of results from regressions (4) and (5), including the full set of control variables. As none of the additional control variables were significant we refrain from commenting on their sign.

population share 60plus on the number of GPs per capita is particularly pronounced for metropolitan districts. This relationship is weakened and eventually overturned for increasingly rural districts. For our preferred specification with interactions, regression (5), the total effect of the share 60plus (0.896-0.999) is now negative in rural districts (type III). The finding that an elderly population becomes progressively less attractive for more rural districts is supported by all regressions with interaction effects. Recall that the coefficient on the share 60plus (including where relevant the interaction terms) can be directly interpreted as the expected profitability of providing services to the elderly as opposed to the middle-aged population from a representative physician's point of view. While the provision of services to a member of the population 60plus is 1.896 times as profitable as the provision to a member of the population 20–59 within a metropolitan district, it is only .897 times as profitable in a rural environment. Thus, our findings lend support to our conjecture that rurality reduces the relative profitability of the elderly population (Hypothesis H2: (a) $\overline{\beta} > 0$. (b) $\overline{\beta}_{III} < \overline{\beta}_{II} < 0$.). Rural areas with their increasing shares of elderly inhabitants are thus exposed to a risk of becoming under-doctored.

When considering all types of districts alike, the share of the young population 19minus has a significant negative effect on the supply of GPs in regressions (4) to (7). This is in line with our Hypothesis H1 ($\underline{\beta} < 0$), according to which a large share of the population below age 20 is relatively unprofitable for GPs, as the young typically turn to paediatricians or specialists. When adding the interaction with rurality, we find that the effect of the share 19minus is more pronounced within rural districts. This may hint at an underlying expectation that young populations will not stay within rural districts but rather migrate elsewhere, implying a reduction in the future demand for GP services.

The growth of average list size within period $t+1$ is significantly positively related to GP supply within period t, an effect we find to be very robust to changes in specification. This indicates that GPs are to some extent anticipating changes in the demand for their services. Thus, we can conclude in line with Hypothesis H3 ($\delta' > 0$) that there is a significant degree of imperfection in equilibrium adjustment. Recall that the coefficient on the growth factor of average list size $\delta' = s\delta$ equals the product of the "average" extent of imperfection $s \in [0,1]$ and the discount factor $\delta = (1+\rho)^{-1}$, where ρ is the discount rate. Using the relationship $s = \delta' (1+\rho)$, with $\delta' \approx 0.481$, in our most preferred specifications (4) and (5), then allows for a back-of-the-envelope assessment of the degree of imperfection in the GP market. Harrison et al. (2002) provide evidence from a field experiment in Denmark suggesting that the discount rate of the "more educated" is 0.206, which would imply $s \approx 0.58$.[47] In a set of additional regressions (not reported here) we have studied the impact of list size growth in periods $t+2$ and $t+3$. The successively smaller coefficients are consistent with the stronger discounting of these more distant effects. There is little impact on the coefficients of the other variables.[48]

The effect of population density is significantly negative and robust to changes in specification. The persistent negative impact of a spatially denser population on GP supply is in line with Hypothesis H4 ($\gamma_1 < 0$), according to which competition is stronger in more densely populated districts. Internists and GPs are substitutes, a result that is robust across different specifications and in line with Hypothesis H5 ($\gamma_2 < 0$). For paediatricians we find a significant positive coefficient, which we interpret as being indicative of the attractiveness of a district in line with Hypothesis H6 ($\gamma_3 > 0$). Regressions (8) to (11) show that the remaining coefficients change only slightly when the two specializations are omitted, except for the share of the population below age 20, which turns insignificant.[49],[50]

[47] A similar result would obtain for the median discount rates 0.275 and 0.3 based on recent experimental evidence for Germany (Dohmen et al. 2010). As one of their key results, the authors show that the discount rate decreases significantly with cognitive ability, implying a somewhat lower value of s. Anderson et al. (2008) demonstrate that estimates of the discount rate are substantially lower when risk aversion is controlled for in a simultaneous estimation. Thus, the value of s reported is likely to be an upper bound.

[48] As list size growth may mask additional and distinct effects of population growth as opposed to the growth in the number of GPs, we have also estimated the model based on the two distinct growth rates. Both variables have the (expected) sign that is consistent with the positive impact of list size growth. All effects are robust to this change in specification.

[49] This is likely to reflect that the variable is now picking up the positive impact of the omitted share of paediatricians as an offsetting force.

[50] In order to rule out any bias from potentially endogenous control variables, e.g. the supply of internists or paediatricians, we have regressed the residuals from specification

For robustness, we consider in Appendix B regressions with the number of SHI-affiliated GPs in a district as dependent variable, treating the size of the population as an independent variable. This specification gives rise to very similar results.

5.3 Accounting for Entry Regulation

We now consider more explicitly the impact of regulatory entry constraints. As outlined earlier, we have no direct information on whether or not entry ceilings were in place within a district in any particular year. We therefore draw on an indicator variable xs_{ijt} which is set to one if and only if district i in state j and year t can be classified as exhibiting "excess supply" according to the regulatory framework. Districts for which an "excess supply" has been established can be subjected to entry ceilings, which in terms of our model imply an upper bound to equilibrium adjustment. Note that whether or not such a ceiling is binding depends on whether the district under consideration still exhibits growing income prospects for physicians (or a growing attractiveness in residential terms). While we cannot observe this, we conjecture that the presence of "excess supply" will be positively correlated with a constraint on future increases in the number of physicians and will, thus, imply that future list size growth is anticipated in the current supply of physicians. Including the interaction term $\delta'' \times xs_{ijt} \times listgro_{ijt+1}$ into (11), we can thus hypothesize that $\delta'' > 0$ (Hypothesis H3').

As described earlier, data limitations restrict us to the construction of either an upper bound measure uxs_{ijt} (counting both GPs and internists against the benchmark) or a lower bound measure lxs_{ijt} (counting only GPs against the benchmark) for the "excess supply" indicator.[51] While we would therefore view our analysis to be merely suggestive of the impacts of entry regulation on physician supply, it nevertheless allows us to establish whether our main results are robust to the explicit incorporation of regulatory effects

(4) on the full set of explanatory variables. The analysis shows that we can rule out any relevant correlation.

[51] Specifically, we set $uxs_{ijt} = 1$ if $GPs\_pc_{ijt} + intern\_pc_{ijt} > 1.1 \times benchmark_{ij}$ and $uxs_{ijt} = 0$ otherwise. Similarly, we set $lxs_{ijt} = 1$ if $GPs\_pc_{ijt} > 1.1 \times benchmark_{ij}$ and $lxs_{ijt} = 0$ otherwise.

into the model.

Table 4 provides the results for the upper and lower bound measures uxs_{ijt} and lxs_{ijt}, respectively, for the two specifications with and without the interaction between the age shares and the rurality levels. We find the coefficient on the interaction between an "excess supply" indicator and list size growth to be significantly positive throughout. This confirms hypothesis H3' that a growing demand for GP services is anticipated in the current supply of GPs especially in those districts in which excess supply leads to the expectation that entry restrictions are about to be introduced. This notwithstanding, we find a significantly positive impact of list size growth on current supply even for districts in which "excess supply" has not been established. We would expect such a finding for the lower bound measure of excess supply [specifications (12) and (13)], as it fails to establish "excess supply" for a range of districts, which may nevertheless be subjected to entry ceilings. However, the result also holds for the upper bound measure [specifications (14) and (15)], which suggests that regulatory entry ceilings are not the only reasons for imperfect equilibrium adjustment.

Our findings with respect to the other variables of interest show that despite some quantitative changes, none of the coefficients change sign and all of the relevant effects remain highly significant. We can therefore conclude that our main results are robust to the inclusion of the "excess supply" indicator.

Table 4 about here

5.4 Decomposing the Change in the Regional Supply of GPs

We conclude our analysis by considering the extent to which demographic and economic development have contributed to the change in the number of GPs per capita over the period 1997–2008. Using the approximation

$$\frac{n_t^*}{\ell_t} \approx \frac{y_t \left[1 + s\delta \left(1 + g_{t+1}^{\ell/n}\right)\right]}{\overline{v}_t \left[1 - (1-s)\,\delta\,(1 + g^{\overline{v}})\right]},$$

which corresponds to our structural equation (10), we obtain the relationship

$$\frac{n^*_{2008}/\ell_{2008}}{n^*_{1997}/\ell_{1997}} = \frac{y_{2008}}{y_{1997}} \times \frac{1 + s\delta\left(1 + g^{\ell/n}_{2009}\right)}{1 + s\delta\left(1 + g^{\ell/n}_{1998}\right)} \times \frac{\overline{v}_{1997}}{\overline{v}_{2008}}.$$

Thus, the growth factor of the number of GPs per capita between 2008 and 1997 can be decomposed into (a) the growth factor of income earned per resident [first right hand side (RHS) term], (b) a multiplier, representing the relative contribution of the next period growth in list size (second RHS term), and (c) the inverse growth factor of the reservation income (third RHS term). While we can calculate the first two terms based on our empirical results, we will infer the growth factor of reservation income by calculating the relationship

$$\frac{\overline{v}_{2008}}{\overline{v}_{1997}} = \left[\frac{y_{2008}}{y_{1997}} \times \frac{1 + s\delta\left(1 + g^{\ell/n}_{2009}\right)}{1 + s\delta\left(1 + g^{\ell/n}_{1998}\right)}\right] \times \left(\frac{n^*_{2008}/\ell_{2008}}{n^*_{1997}/\ell_{1997}}\right)^{-1}.$$

While the income multiplier in (b) can be directly calculated from the data (see below), we show in Appendix A3 that

$$\begin{aligned}\frac{y_{2008}}{y_{1997}} &= \frac{y\left(\overline{\lambda}_{2008}, \underline{\lambda}_{2008}, \zeta_{2008}, r\right)}{y\left(\overline{\lambda}_{1997}, \underline{\lambda}_{1997}, \zeta_{1997}, r\right)} \\ &= \left(1 + \frac{\overline{\beta}(r)\,\Delta\overline{\lambda}_{08/97} + \underline{\beta}(r)\,\Delta\underline{\lambda}_{08/97}}{1 + \overline{\beta}(r)\,\overline{\lambda}_{1997} + \underline{\beta}(r)\,\underline{\lambda}_{1997}}\right) \\ &\quad \times \prod_h \left(\frac{control_{h2008}}{control_{h1997}}\right)^{\gamma_h}, \qquad (12)\end{aligned}$$

where $\Delta\overline{\lambda}_{08/97} := \overline{\lambda}_{2008} - \overline{\lambda}_{1997}$ and $\Delta\underline{\lambda}_{08/97} := \underline{\lambda}_{2008} - \underline{\lambda}_{1997}$, respectively. Combining the data reported in Table 8 in Appendix C with the relevant coefficients from regression (5) in Table 3, we obtain the results reported in Table 5. In calculating the growth factors we employ only variables for which we have found significant coefficients. Namely, these are the age shares in the population as well as their interaction with district type; population density; the per capita numbers of internists and paediatricians, respectively; and the growth factor of average list size.[52]

Overall, the number of GPs per capita has declined across all district types, the decline being most pronounced in municipalized (type II) districts at a rate of 5.4 percent. At the same time, GP income earned per resident has increased across all district types, the increase being at 1.8 percent in metropolitan (type I) and municipalized (type II) districts, and at 6.4 percent in rural (type III) districts.[53]

Interestingly, the change in the population age structure alone would be supporting a stronger increase in income in the more urbanized districts, amounting to 4.1 percent in type I districts and 2.9 percent in type II districts. In these districts, GPs' earnings increase in particular due to the growing share of the elderly population.[54] As expected, an increasing share of the elderly population lowers income in rural districts, but this effect is over-compensated by the decline in the equally unprofitable share of the young population. Overall, we find that population ageing has, in fact, a positive impact on GPs earnings and cannot therefore be responsible for the decline in the provision of GP services over our period of observation.

While the increase in population density in the metropolitan districts tends to lower physician income by 0.8 percent, the decline in population density in municipalized and rural districts tends to increase income. In rural districts, this effect is particularly strong, and would ceteris paribus lead to an increase in income by 1.7 percent. This indicates that the shift in population density from rural districts to the more urban districts has, indeed, a sizeable impact on the competition between GPs, which is weakened, in particular, within rural districts: A spatially more disperse population tends to stifle the incentive to attract patients through the provision of more intensive services which come at a lower profit margin (e.g. late opening hours).

[52] The inclusion of either of the "excess supply" measures from regressions (13) and (15), respectively, does not greatly change the results from our decomposition analysis.

[53] All of the effects reported in the following can be interpreted as ceteris paribus impacts on the number of GPs per capita. Thus, the income growth in type III districts of 6.4 percent would ceteris paribus lead to an increase in the number of GPs per capita by 6.4 percent.

[54] Specifically, the increase in the share 60plus over the time span 1997–2008 tends to increase income by some 3 percent in type I districts and by some 2 percent in type II districts.

The positive impact of population ageing on GP income is dampened in all districts by the increase in the number of internists per capita. Here, the greater provision of substitute services tends to lower income by around 2 percent in all districts. Changes in the number of paediatricians per capita only have a minor impact on GP's income prospects in all types of districts. The positive impact of income growth over the time span 1997–2008 is moderately amplified by the fact that while average list size was expected to shrink in 1998, it was expected to grow in 2009, this effect adding between 0.4 percent to the GP supply in rural districts and 0.9 percent in municipalized districts.

Based on the overall life-cycle income, the GP supply per capita should have grown by 2.3 percent in metropolitan districts, by 2.7 percent in municipalized districts, and by 6.8 percent in rural districts. The actual decline in the supply of GPs then implies an increase in reservation income by some 6.7 percent in metropolitan districts, by 8.6 percent in municipalized districts, and by 12.1 percent in rural districts. A number of trends are consistent with this finding. First, a comparison of physician income from the provision of SHI services between the years 2009 and 2001 shows that income growth across all specialties (and notably for internists) has outpaced the growth of GPs' income (see Busse and Riesberg 2004 and Busse and Blümel 2014).[55] This does not yet include differential income growth from the provision of services to the population with private health insurance, but we would suspect that this would likely reinforce the increase of GPs' reservation income due to greater income growth in other medical specialties.[56] Second, while Ono et al. (2014) report an income premium of around 11 percent for GPs' rural as opposed to urban practice, this is quite possibly insufficient to compensate for a variety of disamenities within rural settings (Günther et al. 2010 and Ono et al. 2014). These relate to working conditions, such as longer office hours and on-call duties within rural regions, and to the living environment, such as

[55] According to Busse and Blümel (2014, p. 156) around 51 percent of a GP's revenue translates into earnings, with the same ratio being 48.8 percent across all specalities and 46.6 percent for internists. Applying these fractions to the revenue levels reported in Tables 3.8 and 3.9 one can calculate the 2009 income levels. These can be compared in a straightforward way to the 2001 income figures reported in Table 32 in Busse and Riesberg (2004).

[56] Differential income growth is also reflected in the divergent supply trends of GPs as opposed to specialists that we report in Section 3.2.

scarcer career opportunities for the partner and less availability of child care and leisure activities (Günther et al. 2010 and Ono et al. 2014). Based on a stated-preference study carried out amongst a representative set of young physicians at the point of establishing practice, Günther et al. (2010) thus report a compensating income of around 9,000 € per month for taking up practice in a rural as opposed to an urban environment. Evidence provided by Ono et al. (2014) suggest that the preference gradient for urban practice is likely to have increased over time.

Table 5 about here

6 Conclusions

We have examined theoretically and empirically how the geographical distribution of office-based physicians, and in our empirical analysis specifically of GPs, responds to differential population change at the regional level. In doing so, we have identified a number of issues pertinent to the provision of physician services over space and time.

First, our finding that the impact of the population share 60plus on the supply of GPs reverses from positive within metropolitan districts to negative within rural districts is direct evidence for the conjecture that the provision of services to an elderly population may be relatively profitable in urban districts but not in a rural environment. While our data do not allow us to identify the causal mechanism, this evidence is consistent with a health care system in which the higher costs of treating elderly patients within a rural context are not fully compensated by the payment system. As we have argued earlier, this holds true for the German reimbursement scheme for panel doctors before 2009 and thus for the period 1997–2008 covered in our regressions. As expected, the share of the young population turns out to be relatively unprofitable from a GP's perspective in all types of districts. Overall these findings indicate that the positive relationship between physician supply and the share of the elderly population as frequently found in cross-sectional studies should be interpreted with caution. A more complex relationship emerges once due account is taken of unobserved regional heterogeneity, of the full age structure of the population, and of the interaction between geographic and demographic structure as mediated through the payment system.

Second, insofar as physicians base long-term decisions about their location on the (expected) profitability from offering treatments to the resident population, the demographic make-up of a region becomes an important determinant not only of the current income but also of the expected income from the provision of services to the future population. Population change may then trigger "anticipatory" changes in the supply of physicians even before it actually takes effect. In our empirical estimation, this implies that the current number of GPs per capita should increase with the rate of list size growth during the subsequent period. While we identify a statistically significant impact of anticipated list size growth on current supply, its magnitude suggests only limited economic relevance.

In order to explain the limited role of anticipation we refer to our third finding of interest: Expectations about future income play a role as a predictor of the current number of physicians if and only if barriers to entry and/or exit prevent the number of physicians from adjusting to its equilibrium value. Under full equilibrium adjustment, physicians expect that changes in future income will be offset by changes in the number of rival physicians in a way that equilibrates the future income stream with the stream of reservation income. In this case, the current supply of physicians is independent of population change and the evolution of other profit shifters. Thus, we would expect anticipatory effects predominantly in regions which are subject to entry ceilings, and we can confirm this empirically: The impact of list size growth is significantly larger in regions that are classified as exhibiting an "excess supply" according to regulatory rules and therefore more likely to be subject to entry controls.

Fourth, when decomposing the change in the per-capita supply of GPs over the time span 1997–2008 we find, perhaps surprisingly, that population ageing is contributing to an increase in physician income in all types of districts alike. Within metropolitan and urban districts, GPs benefit both from an increase in the share of the elderly population as well as from the reduction in the share of the young population While the increase in the share of the elderly population is lowering GPs' earnings in rural districts, this effect is compensated by the strong decline in the young population. An additional boost to GPs' earnings in rural districts arises from the fall in population density, which can be argued to lead to less intensive competition. The increase in physician income that is implied by our decomposition analysis should ceteris paribus lead to an increase in the number of GPs per capita in all types of districts. The fact that we observe a decline should therefore be explained by an increasing reservation income of becoming a GP, which is most pronounced in rural areas.

For policy-makers interested in ensuring an adequate and equitable provision of health services across regions, it is important to know whether the supply of physicians adjusts appropriately to changes in the local demand for services. Given that population ageing is widely expected to come with an increased per-capita demand for ambulatory health care, our results carry a mixed message for Germany: While an increase in the share of the elderly population is ceteris paribus prone to induce a greater supply of GPs in metropolitan districts, the opposite is true for rural districts. This suggests that within rural areas the local supply of physicians may not be adjusting to the needs of an ageing population. The divergent impact across the urban–rural spectrum also suggests that the gap between an over-provision of physicians within metropolitan areas and an under-provision within rural areas may well be widening. For GPs, this tendency is offset, however, by the decline in the share of the young population, implying that the overall development will strongly depend on the particular pattern of the ageing dynamics. The supply of GP services is further shaped by a shift in the concentration of the population away from rural regions into metropolitan and urban regions: the increase in population density within metropolitan areas tends to depress the number of GPs per capita, whereas the decline in population density within rural regions tends to cushion the fall in the per-capita supply of GPs. While this dampens to some extent the widening in the urban–rural gap in the per-capita supply of GPs, the underlying change in competition is prone to imply a greater provision *at the level of the individual physician* within urban areas and a lower provision within rural areas. Finally, our analysis shows that for the period covered by our data, the population-driven trends in physician supply are strongly overridden by an opportunity- and preference-driven downward trend in the supply of GPs, especially in rural areas. Whether or not this continues to be true in the light of continued ageing remains to be assessed by future research.

Altogether, our findings tie in with general concerns about regional inequality in the provision of health care in Germany (Fülöp et al. 2010, Advisory Council 2014, Ozegowski and Sundmacher 2014) as well as in other OECD countries (Cooper et al. 2002, Arcury et al. 2005, Morris et al. 2005, Goddard et al. 2010, Ono et al. 2014). Various policy initiatives are on the way in Germany (Advisory Council 2014, Ozegowski and Sundmacher 2014) and elsewhere (Ono et al. 2014) to encourage physicians, and especially GPs, to take up practice in rural areas or to offset a shortfall by different arrangements of provision, including the substitution of nurse practitioners and assistants for physicians (Stange 2014). While an evaluation of these policy measures lies outside the scope of this paper, our results suggest that significant adjustments to reimbursement may be necessary in order to stabilize the supply of GPs in rural areas in Germany. The level of remuneration of GPs needs to be increased substantially in order to match the income growth for specialists, and substantial income premia are required to stimulate entry into rural practice. Indeed, both of these factors are recognized in recent policy proposals (Advisory Council 2014, Ozegowski and Sundmacher 2014). Accompanying reforms are likely to be necessary in respect to the mode of provision. Here, the strengthening of group practices and primary care centers may both facilitate a continuous provision of care and render the working environment more attractive, especially for young physicians who exhibit a lesser preference for single-handed practice (Ono et al. 2014).

Our results also yield a couple of general insights. First, the financial incentives that come through the reimbursement scheme should not be based on geographic or demographic criteria in isolation when it comes to an equitable provision of services to all population groups in all regions. In case of Germany, for instance, the negative urban-to-rural gradient in the relative profitability of the population share 60plus suggests that physician reimbursement should be adjusted according to both age and regional character rather than according to one criterion alone. The 2007 and 2011 health care reforms have to some extent gone down this way by introducing adjustments in the fees to the morbidity structure and to (exceptional) regional circumstances as well as by granting greater regional autonomy in the design of payments (Busse and Blümel 2014). Second, policy initiatives should not only embrace short-term incentives for physicians to settle in under-doctored areas, but also the economic and demographic prospects which are prone to shape the supply of physicians in the longer term. Indeed, regional population change poses a multi-faceted challenge for health policy: population structure interacts with geographic structure both in a static way and over time, and this interaction is strongly shaped by the reimbursement scheme. This opens considerable scope for future work both on the extent to which alternative (e.g. more capitation-oriented) reimbursement schemes are conducive (or not) to a provision of physician services that is equitable across sub-populations, space and time, as well as on the characterization of an "optimal" scheme in that respect.

7 References

Advisory Council on the Assessment of Developments in the Health Care System, 2014, Needs-based Health Care: Opportunities for Rural Regions and Selected Health Care Sectors. Report 2014.

Andersen, S., Harrison, G.W., Lau, M.I., Rutström, E.E., 2008, Eliciting risk and time preferences, *Econometrica* 76, 583-618.

Arcury, T.A., Gesler, W.M., Preisser, J.S., Sherman, J., Spencer, J., Perin, J., 2005, The Effects of Geography and Spatial Behavior on Health Care Utilization among the Residents of a Rural Region, *Health Services Research* 40, 135-155.

Aylin, P., Majeed, F.A., Cook, D.G., 1996, Home Visiting by General Practitioners in England and Wales, *British Medical Journal* 313, 207-210.

Benstetter, F., Wambach, A., 2006, The treadmill effect in a fixed budget system, *Journal of Health Economics* 25, 146-169.

Boerma, W.G.W., Groenewegen, P.P., 2001, GP Home Visiting in 18 European Countries. Adding the Role of Health System Features, *European Journal of General Practice* 7, 132-137.

Bolduc, D., Fortin, B., Fournier, M.-A., 1996, The Effect of Incentive Policies on the Practice Location of Doctors: A Multinomial Probit Analysis, *Journal of Labor Economics* 14, 703-723.

Bresnahan, T.F., Reiss, P.C., 1991, Entry and Competition in Concentrated Markets, *Journal of Political Economy* 99, 977-1009.

Breyer, F., Mühlenkamp, H., Adam, H., 1986, Determinants of the utilization of physician services in the system of statutory health insurance in Germany, in: G.v.d. Schulenburg, J.-M. (ed.) Essays in social security economics, Berlin: Springer.

Busse, R., Blümel, B., 2014, Germany: Health System Review, *Health Care Systems in Transition* 16, 1-296.

Busse, R., Riesberg, A., 2004, Health Care Systems in Transition: Germany, Copenhagen: WHO Regional Office on Behalf of the European Observatory on Health Systems and Policies.

Chaix, B., Veugelers, P.J., Boelle, P.-Y., Chauvin, P., 2005, Access to General Practitioner Services: the Disabled Elderly Lag Behind in Underserved Areas, *European Journal of Public Health* 15, 282-287.

Colwill, J.M., Cultice, J.M., Kruse, R.L., 2008, Will Generalist Physician Supply Meet Demands of an Increasing and Aging Population, *Health Affairs* 27, w232-w241.

Cooper, R.A. et al., 2002, Economic and Demographic Trends Signal an Impending Physician Shortage, *Health Affairs* 21, 140-154.

Correia, I., Veiga, P., 2010, Geographic distribution of physicians in Portugal, *European Journal of Health Economics* 11, 383-393.

Dionne, G., Langlois, A., Lemire, N., 1987, More on the geographical distribution of physicians, *Journal of Health Economics* 6, 365-374.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., 2010, Are risk aversion and impatience related to cognitive ability? *American Economic Review* 100, 1238-1260.

Dormont, B., Samson, A.-L., 2008, Medical Demography and Intergenerational Inequalities in General Practitioner's Earnings, *Health Economics* 17, 1037-1055.

Driscoll, J.C., Kraay, A.C., 1998, Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data, *Review of Economics and Statistics* 80, 549-560.

Dusheiko, M, Gravelle, H., Campbell, S., 2002, Inequality in Consultations with General Practitioners, CHE Technical Paper 24.

Elliott, B., et al., 2006, Adjusting the General Medical Services Allocation Formula to Reflect Recruitment and Retention Difficulties, Final Report, University of Aberdeen: Health Economics Research Unit.

Farbmacher, H., Winter J., 2013, Per-period co-payments and the demand for health care: evidence from survey and claims data, *Health Economics* 22, 1111-1123.

Foster, S.A., Gorr, W.L., 1992, Federal Health Care Policy and the Geographic Diffusion of Physicians: A Macro-Scale Analysis, *Policy Sciences* 25, 117-134.

Fülöp, G., Kopetsch, T., Hofstätter, G., Schöpe, P., 2008, Regional distribution effects of "needs planning" for office-based physicians in Germany and Austria – methods and empirical findings, *Journal of Public Health* 16, 447-455.

Fülöp, G., Kopetsch, T., Schöpe, P., 2010, Planning medical care for actual need, *Journal of Public Health* 18, 97-104.

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, *Annals of Regional Science* 46, 691-706.

Gächter, M., Schwazer, P., Theurl, E., Winner, H., 2012, Physician Density in a Two-tiered Health Care System, *Health Policy* 106, 257-268.

Gerste, B., 2012, Die Inanspruchnahme von Gesundheitsleistungen im Alter [Health care utlization in old age], in: Günster, C., Klose, J., Schmacke, N., (eds.). Versorgungs-Report 2012, Stuttgart: Schattauer.

Giuffrida, A., Gravelle, H., 2001, Inducing or Restraining Demand: The Market for Night Visits in Primary Care, *Journal of Health Economics* 20, 755-779.

Goddard, M., Gravelle, H., Hole, A., Marini, G., 2010, Where did all the GPs go? Increasing supply and geographical equity in England and Scotland, *Journal of Health Service Research and Policy* 15, 28-35.

Gravelle, H., 1999, Capitation contracts: access and quality, *Journal of Health Economics* 18, 315-340.

Günther, O.H., Kürstein, B., Riedel-Heller, S.G., König, H.-H., 2010, The role of monetary and nonmonetary incentives on the choice of practice establishment: a stated preference study of young physicians in Germany, *Health Services Research* 45, 212-229.

Hann, M., Gravelle, H., 2004, The Maldistribution of General Practitioners in England and Wales: 1974-2003, *British Journal of General Practice* 54, 894-898.

Harrison, G.W., Lau, M.I., Williams, M.B., 2002, Estimating individual discount rates for Denmark: a field experiment, *American Economic Review* 92, 1606-1617

Hingstman, L., Boon, H., 1989, Regional Dispersion of Independent Professionals in Primary Health Care in the Netherlands, *Social Science & Medicine* 28, 121-129.

Hurley, J.E., 1991, Physicians' Choices of Specialty, Location and Mode: A Reexamination within an Interdependent Decision Framework, *Journal of Human Resources* 26, 47-71.

Iversen, T., Kopperud, G.S., 2005, Regulation versus Practice - the Impact of Accessibility on the Use of Specialist Health Care in Norway, *Health Economics* 14, 1231-1238.

Jürges, H., 2007, Health Insurance Status and Physician-Induced Demand for Medical Services in Germany: New Evidence from Combined District and Individual Level Data, DIW Discussion Paper 689.

Jürges, H., 2009, Health Insurance Status and Physician Behavior in Germany, *Schmollers Jahrbuch* 129, 297-307.

Kamtsiuris, P., Bergmann, E., Rattay, P., Schlaud, M., 2007, Inanspruchnahme medizinischer Leistungen: Ergebnisse des Kinder- und Jugendgesundheitssurveys (KiGGS) [Health care utilization: results from the survey on child and adolescent health], *Bundesgesundheitsblatt* 50, 836-850.

Klose, J., Rehbein, I., 2011, Ärzteatlas 2011: Daten zur Versorgungsdichte von Vertragsärzten [Physician Atlas 2011: data on the supply of panel doctors], Wissenschaftliches Institut der AOK.

Kopetsch, T., 2007, Arztdichte und Inanspruchnahme ärztlicher Leistungen in Deutschland. Eine empirische Untersuchung der These von der angebotsinduzierten Nachfrage nach ambulanten Arztleistungen [Physician density and use of physician services in Germany. An empirical investigation into the thesis of supplier-induced demand], *Schmollers Jahrbuch* 127, 373-405.

Kopetsch, T., Munz, H., 2007, Analyse des Niederlassungs- und Wanderungsverhaltens von Vertragsärzten und -psychotherapeuten in Deutschland [Analysis of the Location and Migration Choices by SHI Contract Physicians in Germany], Berlin: Kassenaerztliche Bundesvereinigung.

Kopetsch, T., Schmitz, H., 2014, Regional variation in the utilisation of ambulatory services in Germany, *Health Economics* 23, 1481-1492.

Kraft, K., von der Schulenburg, J.-M., 1986, Co-Insurance and Supplier-Induced Demand in Medical Care: What Do We Have to Expect as the

Physician's Response to Increased Out-of-Pocket Payments, *Journal of Institutional and Theoretical Economics* 142, 360-379.

Marshall, T., 1996, Home Visits are a Feature of Primary Care in Many Western European Countries, *British Medical Journal* 313, 1085.

Meyer, G.S., Gibbons, R.V., 1997, House Calls to the Elderly – A Vanishing Practice Amongst Physicians, *New England Journal of Medicine* 337, 1815-1820.

Morris, S., Sutton, M., Gravelle, H., 2005, Inequity and Inequality in the Use of Health Care in England: an Empirical Investigation, *Social Science & Medicine* 60, 1251-1266.

Newhouse, J.P., 1990, Geographic Access to Physician Services, *Annual Review of Public Health* 11, 207-230.

Newhouse, J.P., Williams, A.P., Bennett, B.W., Schwartz, W.B., 1982, Does the geographical distribution of physicians reflect market failure?, *Bell Journal of Economics* 13, 493-505.

Nocera, S., Wanzenried, G., 2008, On the Dynamics of Physician Density: Theory and Empirical Evidence from Switzerland, *Swiss Journal of Economics and Statistics* 104, 247-282.

Nuscheler, R., 2003, Physician Reimbursement, Time Consistency, and the Quality of Care, *Journal of Institutional and Theoretical Economics* 159, 302-322.

Ono, T., Schoenstein, M., Buchan, J., 2014, Geographic Imbalances in Doctor Supply and Policy Responses, OECD Health Working Paper No. 69.

Ozegowski, S., Sundmacher, L., 2012, Ensuring access to health care – Germany reforms supply structures to tackle inequalities, *Health Policy* 106, 105-109.

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.

Robert Koch-Institute (RKI), 2008, Health in Germany, Federal Health Reporting. Berlin.

Rosenthal, M.B., Zaslavsky, A., Newhouse, J.P., 2005, The Geographic Distribution of Physicians Revisited, *Health Services Research* 40, 1931-1952.

Salop, S.C., 1979, Monopolistic competition with outside goods, *Bell Journal of Economics* 10, 141-156.

Schaumans, C., Verboven, F., 2008, Entry and Regulation: Evidence from Health Care Professions, *RAND Journal of Economics* 39, 949-972.

Schmitz, H., 2013, Practice budgets and the patient mix of physicians – The effect of a remuneration system reform on health care utilisation, *Journal of Health Economics* 32, 1240-1249.

Scholz, S., Graf von der Schulenburg, J.-M., Greiner, W., 2015, Regional differences of outpatient physician supply as theoretical economic and empirical generalized linear model, *Human Resources for Health* 13.

Schreyögg, J., Grabka, M.M., 2010, Copayments for ambulatory care in Germany: a natural experiment using a diference-in-difference approach, *European Journal of Health Economics* 11, 331-341.

Sibbald, B., Bojke, C., Gravelle, H., 2003, National survey of job satisfaction and retirement intentions among general practitioners in England, *British Medical Journal* 326.

Stange, K., 2014, How does Provider Supply and Regulation Influence Health Care Markets? Evidence from Nurse Praticitioners and Physician Assistants, *Journal of Health Economics* 33, 1-27.

Straand, J., Sandvik, H., 1998, Home Visits in General Practice – Most Often for Elderly Patients: A Report from the Möre & Romsdal Prescription Study, *Norwegian Journal of Epidemiology* 8, 127-132.

Sundmacher, L., Busse, R., 2011, The impact of physician supply on avoidable cancer mortality in Germany, *Health Policy* 203, 53-62.

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.

Sundmacher, L., Ozegowski, S., 2016, Regional distribution of physicians: the role of comprehensive private health insurance in Germany, *European Journal of Health Economics* 17, 443-451.

Taylor, D.H., Leese, B., 1998, General practitioner turnover and migration in England 1990-94, *British Journal of General Practice* 48, 1070-1072.

Thode, N., Bergmann, E., Kamtsiuris, P., Kurth, B.-M., 2005, Einflussfaktoren auf die ambulante Inanspruchnahme in Deutschland [Determinants of the usage of ambulatory care in Germany], *Bundesgesundheitsblatt* 48, 296-306.

Vogt, V., 2016, The contribution of locational factors to regional variations in office-based physicians in Germany, *Health Policy* 120, 198-204.

8 Appendix A

8.1 Appendix A1: General Model and Proof of Proposition 1

We derive the equilibrium for a general multi-period model of regional physician entry and show that the results presented in Proposition 1 for the two-period case are entirely general. Consider a set-up where a representative physician practises for $1 + z$, with $z \geq 1$, periods after which she retires or leaves the region/profession for other reasons. We focus on the physician's life-cycle problem within a single region. While we assume that an entry equilibrium is realized in period t, the representative physician expects for all subsequent periods $\widehat{t} \in [t+1, t+z]$ that physician supply does not fully adjust to equilibrium with probability $s \in [0,1]$. Let

$$E\phi_{\widehat{t}} := y_{\widehat{t}} \left[s\frac{\ell_{\widehat{t}}}{\widehat{n}_{\widehat{t}}} + (1-s)\,\frac{\ell_{\widehat{t}}}{n^*_{\widehat{t}}} \right]$$

then denote the expected income for period $\widehat{t}$, where $y_{\widehat{t}} = y\left(\overline{\lambda}_{\widehat{t}}, \underline{\lambda}_{\widehat{t}}, \zeta_{\widehat{t}}, r\right)$, as given by (1), denotes the average expected income per resident at time $\widehat{t}$, and where $\ell_{\widehat{t}}/n_{\widehat{t}}$ and $\ell_{\widehat{t}}/n^*_{\widehat{t}}$ denote the list sizes corresponding to a disequilibrium number of physicians $\widehat{n}_{\widehat{t}}$ and an equilibrium number $n^*_{\widehat{t}}$, respectively. Applying a discount factor $\delta < 1$, we can write the present value of the physician's expected life-cycle income at time t as

$$v_t = \sum_{\widehat{t}=t}^{t+z} \delta^{\widehat{t}-t} E\phi_{\widehat{t}}. \tag{13}$$

We continue to employ the utility function (2), where the benefit from regional amenities is now given by $u_t = \sum_{\widehat{t}=t}^{t+z} \delta^{\widehat{t}-t} u_{\widehat{t}}$. Continuing to describe the reservation income by (3) and defining $\mathbf{En}_{t,z} := \{En_{\widehat{t}}\}_{\widehat{t}=t+1}^{t+z}$ as the set of the expected number of physicians for the periods $t+1$ through $t+z$, i.e. up to retirement, an entry equilibrium is given by the number of physicians n^*_t that satisfies $v_t\left(n^*_t, \mathbf{En}_{t,z}\right) = \overline{v}_t$. Observe that in any period $\widehat{t} \in [t+1, t+z]$ the physician supply has been in continuous disequilibrium (since the last equilibrium point in t) with probability $s^{\widehat{t}-t}$ and express the disequilibrium supply in period $\widehat{t}$ as $n_{\widehat{t}} = \psi_{\widehat{t}} n^*_t$, with

$$\psi_{\widehat{t}} \in \begin{cases} \left[1, n^*_{\widehat{t}}/n^*_t\right) & \textit{for } n^*_{\widehat{t}}/n^*_t > 1 \\ \left(n^*_{\widehat{t}}/n^*_t, 1\right] & \textit{for } n^*_{\widehat{t}}/n^*_t < 1 \end{cases} \tag{14}$$

denoting the extent of (partial) adjustment towards the equilibrium value $n^*_{\widehat{t}}$. We can now derive the following result.

Proposition 1A *An entry equilibrium at time t with $s \in [0,1]$ and $z \geq 1$*

sufficiently large, is (approximately) given by

$$\frac{n_t^*}{\ell_t} \approx \frac{y_t + \sum_{\widehat{t}=t+1}^{t+z} (s\delta)^{\widehat{t}-t} \frac{y_{\widehat{t}}\ell_{\widehat{t}}}{\psi_{\widehat{t}}\ell_t}}{\overline{v}_t - (1-s)\,\delta \sum_{\widehat{t}=t+1}^{t+z} (s\delta)^{\widehat{t}-t-1}\,\overline{v}_{\widehat{t}}}. \tag{15}$$

Proof: Consider an entry equilibrium in period t and expand the expression of a physician's life-cycle income as follows:

$$\begin{aligned}
\overline{v}_t &= v_t\left(n_t^*, \mathbf{En}_{t,z}\right) = \frac{y_t\ell_t}{n_t^*} + \sum_{\widehat{t}=t+1}^{t+z} \delta^{\widehat{t}-t} E\phi_{\widehat{t}} \\
&= \frac{y_t\ell_t}{n_t^*} + (1-s)\,\delta\left(\frac{y_{t+1}\ell_{t+1}}{n_{t+1}^*} + \sum_{\widehat{t}=t+2}^{t+z} \delta^{\widehat{t}-t} E\phi_{\widehat{t}}\right) + s\delta\left(\frac{y_{t+1}\ell_{t+1}}{\psi_{t+1}n_t^*} + \sum_{\widehat{t}=t+2}^{t+z} \delta^{\widehat{t}-t} E\phi_{\widehat{t}}\right) \\
&= \frac{y_t\ell_t}{n_t^*} + (1-s)\,\delta\left(\frac{y_{t+1}\ell_{t+1}}{n_{t+1}^*} + \sum_{\widehat{t}=t+2}^{t+z} \delta^{\widehat{t}-t} E\phi_{\widehat{t}}\right) \\
&\quad + s\delta\left[\frac{y_{t+1}\ell_{t+1}}{\psi_{t+1}n_t^*} + (1-s)\,\delta\left(\frac{y_{t+2}\ell_{t+2}}{n_{t+2}^*} + \sum_{\widehat{t}=t+3}^{t+z} \delta^{\widehat{t}-t} E\phi_{\widehat{t}}\right) + s\delta\left(\frac{y_{t+2}\ell_{t+2}}{\psi_{t+1}n_t^*} + \sum_{\widehat{t}=t+3}^{t+z} \delta^{\widehat{t}-t} E\phi_{\widehat{t}}\right)\right] \\
&= \ldots \\
&= \frac{y_t\ell_t}{n_t^*} + \sum_{\widehat{t}=t+1}^{t+z} (s\delta)^{\widehat{t}-t} \frac{y_{\widehat{t}}\ell_{\widehat{t}}}{\psi_{\widehat{t}}n_t^*} + (1-s)\,\delta \sum_{\widehat{t}=t+1}^{t+z} (s\delta)^{\widehat{t}-t-1} \left(\frac{y_{\widehat{t}}\ell_{\widehat{t}}}{n_{\widehat{t}}^*} + \sum_{\widehat{\widehat{t}}=\widehat{t}+1}^{t+z} \delta^{\widehat{\widehat{t}}-\widehat{t}} E\phi_{\widehat{\widehat{t}}}\right),
\end{aligned} \tag{16}$$

where the last equality follows from a collection of all terms relating to continuing disequilibrium (the first sum from $t+1$) and to a switch to equilibrium (the second sum from $t+1$). By definition of an entry equilibrium we have

$$\frac{y_{\widehat{t}}\ell_{\widehat{t}}}{n_{\widehat{t}}^*} = \overline{v}_{\widehat{t}} - \sum_{\widehat{\widehat{t}}=\widehat{t}+1}^{\widehat{t}+z} \delta^{\widehat{\widehat{t}}-\widehat{t}} E\phi_{\widehat{\widehat{t}}}.$$

Inserting this into (16), we obtain

$$\begin{aligned}
\overline{v}_t &= v_t\left(n_t^*, \mathbf{En}_{t,z}\right) = \frac{y_t \ell_t}{n_t^*} + \sum_{\widehat{t}=t+1}^{t+z} (s\delta)^{\widehat{t}-t} \frac{y_{\widehat{t}} \ell_{\widehat{t}}}{\psi_{\widehat{t}} n_t^*} \\
&\quad + (1-s)\,\delta \sum_{\widehat{t}=t+1}^{t+z} (s\delta)^{\widehat{t}-t-1} \left(\overline{v}_{\widehat{t}} - \sum_{\widehat{\widehat{t}}=t+z+1}^{\widehat{t}+z} \delta^{\widehat{\widehat{t}}-\widehat{t}} E\phi_{\widehat{\widehat{t}}} \right) \\
&\approx \frac{y_t \ell_t}{n_t^*} + \sum_{\widehat{t}=t+1}^{t+z} (s\delta)^{\widehat{t}-t} \frac{y_{\widehat{t}} \ell_{\widehat{t}}}{\psi_{\widehat{t}} n_t^*} + (1-s)\,\delta \sum_{\widehat{t}=t+1}^{t+z} (s\delta)^{\widehat{t}-t-1} \overline{v}_{\widehat{t}}, \qquad (17)
\end{aligned}$$

where the approximation follows when noting that $\sum_{\widehat{t}=t+1}^{t+z} (s\delta)^{\widehat{t}-t-1} \sum_{\widehat{\widehat{t}}=t+z+1}^{\widehat{t}+z} \delta^{\widehat{\widehat{t}}-\widehat{t}} E\phi_{\widehat{\widehat{t}}} = \delta^z \sum_{\widehat{t}=z+1}^{2z} \frac{s^z - s^{\widehat{t}-z-1}}{s-1} \delta^{\widehat{t}-z-1} E\phi_{t+\widehat{t}}$ approaches zero for $\delta < 1$ and z sufficiently large. Solving the expression in (17) for n_t^*/ℓ_t yields (15).■

The expression reported in (5) of part (i) of Proposition 1 then follows immediately when setting $z = 1$ in (15). The interpretation of the more general result is analogous to the one presented for the two-period case. According to (15), the equilibrium number of physicians per capita at time t increases with the current income, y_t, as well as with the discounted stream of income over the physician's remaining working life that is expected for a continuing disequilibrium configuration from period $t+1$ onward. This stream increases with the extent of (disequilibrium) list-size growth. Notably, the discount factor and the probability of continuing disequilibrium are compounded, implying that future income streams would typically receive low weights.[57] Conversely, the equilibrium number of physicians per capita decreases with the current reservation income $\overline{v}_t$ and increases with the discounted stream of reservation income over the remaining working life that is expected for a reswitching of the market into equilibrium. Again, the reservation income in more distant periods receives low weights due to the expectation that a reswitching of the

[57] Of course, there may be several spells of disequilibrium over a physician's working life. Note, however, that any intermittent equilibrium in a period $\widehat{t} \in [t+1, t+z]$ will lead to a replacement of the income stream over the time span $\left[\widehat{t}, t+z\right]$ with the reservation income $\overline{v}_{\widehat{t}}$ less an adjustment term for the time span $\left[t+z+1, \widehat{t}+z\right]$, which in the presence of discounting is negligible for z sufficiently large. Hence, conditional on an equilibrium in period $\widehat{t}$ disequilibrium configurations for $\widehat{\widehat{t}} \in \left[\widehat{t}, t+z\right]$ have no bearing on the supply of physicians in period t.

market would have occurred much earlier. Finally, note that the boundary cases $s = 0$ and $s = 1$ follow immediately along the lines discussed in the context of Proposition 1.

Part (ii) of Proposition 1 can be proved as follows: Evaluating (5) at $\psi_{t+1} = n^*_{t+1}/n^*_t$ and re-solving the resulting expression yields

$$\frac{n^*_t}{\ell_t}\bigg|_{s\geq 0,\psi_{t+1}=n^*_{t+1}/n^*_t} = \frac{y_t}{\overline{v}_t - (1-s)\,\delta\overline{v}_{t+1} - s\delta y_{t+1}\frac{\ell_{t+1}}{n^*_{t+1}}} \approx \frac{y_t}{\overline{v}_t - \delta\overline{v}_{t+1}} = \frac{n^*_t}{\ell_t}\,|_{s=0}\ ,$$

where the approximation follows when observing that $\delta y_{t+1}\frac{\ell_{t+1}}{n^*_{t+1}} = \delta\overline{v}_{t+1} + \delta^2 E\phi_{t+2} \approx \delta\overline{v}_{t+1}$ under the premise that physicians will not give much consideration to the income/market in period $t+2$ after their retirement.[58]

Furthermore, it is easy to check that $sgn\frac{\partial}{\partial s}\left(\frac{n^*_t}{\ell_t}\right) = sgn\Omega\left(\psi_{t+1}\right)$ with

$$\Omega\left(\psi_{t+1}\right) = \frac{y_{t+1}\ell_{t+1}}{\psi_{t+1}\ell_t}\left(\overline{v}_t - \delta\overline{v}_{t+1}\right) - y_t\overline{v}_{t+1}.$$

Note that $\Omega' < 0$ and

$$\Omega\left(\frac{n^*_{t+1}}{n^*_t}\right) = \frac{n^*_t}{\ell_t}\frac{y_{t+1}\ell_{t+1}}{n^*_{t+1}}\left(\overline{v}_t - \delta\overline{v}_{t+1}\right) - y_t\overline{v}_{t+1} = \frac{y_{t+1}\ell_{t+1}}{n^*_{t+1}}y_t - y_t\overline{v}_{t+1} \approx 0,$$

where the second equality follows when inserting $\frac{n^*_t}{\ell_t}\Big|_{s\geq 0,\psi_{t+1}=n^*_{t+1}/n^*_t} = \frac{y_t}{\overline{v}_t-\delta\overline{v}_{t+1}}$ and where the approximation follows again under the presumption that $\delta^2 E\phi_{t+2} \to 0$. Observing that $n_{t+1} = \psi_{t+1}n^*_t \lesseqgtr n^*_{t+1}$ implies $\psi_{t+1} \lesseqgtr n^*_{t+1}/n^*_t$, we obtain the result reported.

[58] Note that for the two period case, in which each period spans about 20 years, we will have $\delta << 1$ such that $\delta^2 \to 0$ is a reasonable assumption.

8.2 Appendix A2: Derivation of structural equation (10)

Taking logs of the expression in (7) gives

$$\ln\left(\frac{n^*_{it}}{\ell_{it}}\right) = \ln y_{it} + \ln\left[1 + s\delta\left(1 + g^y_{it+1}\right)\left(1 + g^{\ell/n}_{it+1}\right)\right] - \ln \overline{v}_{it} - \ln\left[1 - (1-s)\,\delta\left(1 + g^{\overline{v}}_{it+1}\right)\right]. \tag{18}$$

We can now express the first summand on the RHS as

$$\begin{aligned} \ln y_{it} &= \ln\left\langle\left\{y(2,r_i) + \left[y(3,r_i) - y(2,r_i)\right]\overline{\lambda}_{it} + \left[y(1,r_i) - y(2,r_i)\right]\underline{\lambda}_{it}\right\}\zeta_{it}\right\rangle \\ &= \ln y(2,r_i) + \ln\left[1 + \overline{\beta}(r_i)\,\overline{\lambda}_{it} + \underline{\beta}(r_i)\,\underline{\lambda}_{it}\right] + \ln\zeta_{it} \\ &\approx \ln y(2,r_i) + \overline{\beta}(r_i)\,\overline{\lambda}_{it} + \underline{\beta}(r_i)\,\underline{\lambda}_{it} + \ln\zeta_{it}, \end{aligned} \tag{19}$$

with $\overline{\beta}(r_i)$ and$\underline{\beta}(r_i)$ as defined in (8) and (9), respectively. Assuming (A7) $\overline{\beta}(r_i)\,\overline{\lambda}_{it} + \underline{\beta}(r_i)\,\underline{\lambda}_{it} \ll 1$, the third line then follows from a first-order Taylor approximation.[59]

The second summand on the RHS of (18) can be expressed as

$$\begin{aligned} &\ln\left[1 + s\delta\left(1 + g^y_{it+1}\right)\left(1 + g^{\ell/n}_{it+1}\right)\right] \\ \approx\;& s\delta\left(1 + g^y_{it+1}\right)\left(1 + g^{\ell/n}_{it+1}\right) \approx s\delta\left(1 + g^{\ell/n}_{it+1}\right), \end{aligned} \tag{20}$$

where the first approximation follows from a first-order Taylor approximation when assuming $\delta \ll s^{-1}$ as in (A2') and where the second approximation follows when assuming $g^y_{it+1} \to 0$ as in (A4).

The third and fourth summands on the RHS of (18) can be expressed as

$$\begin{aligned} &-\ln\overline{v}_{it} - \ln\left[1 - (1-s)\,\delta\left(1 + g^{\overline{v}}_{it+1}\right)\right] \\ =\;& -\ln\overline{w}_t + \ln u_i - \ln\left[1 - (1-s)\,\delta\left(1 + g^{\overline{v}}\right)\right], \end{aligned} \tag{21}$$

where the equality follows when observing (3) and assumptions (A5) and (A6). Inserting (19)–(21) into (18), taking account of (A8), summarizing

[59] The term $\ln\left[1 + \overline{\beta}(r_i)\,\overline{\lambda}_{it} + \underline{\beta}(r_i)\,\underline{\lambda}_{it}\right]$ is defined only for $\overline{\beta}(r_i)\,\overline{\lambda}_{it} + \underline{\beta}(r_i)\,\underline{\lambda}_{it} \geq -1$. This must necessarily be true. Noting (A1), it is readily verified from (8) and (9) that $\min\left\{\overline{\beta}(r_i), \underline{\beta}(r_i)\right\} \geq -1$. But then $\overline{\beta}(r_i)\,\overline{\lambda}_{it} + \underline{\beta}(r_i)\,\underline{\lambda}_{it} \geq -\left(\overline{\lambda}_{it} + \underline{\lambda}_{it}\right) \geq -1$.

terms and assigning parameters, we obtain the structural equation (10) with $\delta' = s\delta$.

8.3 Appendix A3: Derivation of equation (12)

It is easily verified that

$$\begin{aligned}
\frac{y_{2008}}{y_{1997}} &= \frac{y\left(\overline{\lambda}_{2008}, \underline{\lambda}_{2008}, \zeta_{2008}, r\right)}{y\left(\overline{\lambda}_{1997}, \underline{\lambda}_{1997}, \zeta_{1997}, r\right)} \\
&= \frac{y(2,r) + [y(3,r) - y(2,r)]\,\overline{\lambda}_{2008} + [y(1,r) - y(2,r)]\,\underline{\lambda}_{2008}}{y(2,r) + [y(3,r) - y(2,r)]\,\overline{\lambda}_{1997} + [y(1,r) - y(2,r)]\,\underline{\lambda}_{1997}} \times \frac{\zeta_{2008}}{\zeta_{1997}} \\
&= \left\{1 + \frac{[y(3,r) - y(2,r)]\,\Delta\overline{\lambda}_{08/97} + [y(1,r) - y(2,r)]\,\Delta\underline{\lambda}_{08/97}}{y(2,r) + [y(3,r) - y(2,r)]\,\overline{\lambda}_{1997} + [y(1,r) - y(2,r)]\,\underline{\lambda}_{1997}}\right\} \times \frac{\zeta_{2008}}{\zeta_{1997}} \\
&= \left(1 + \frac{\overline{\beta}(r)\,\Delta\overline{\lambda}_{08/97} + \underline{\beta}(r)\,\Delta\underline{\lambda}_{08/97}}{1 + \overline{\beta}(r)\,\overline{\lambda}_{1997} + \underline{\beta}(r)\,\underline{\lambda}_{1997}}\right) \times \prod_h \left(\frac{control_{h2008}}{control_{h1997}}\right)^{\gamma_h},
\end{aligned}$$

where the second equality follows when inserting from (1); where the third equality follows when substituting $\Delta\overline{\lambda}_{08/97} := \overline{\lambda}_{2008} - \overline{\lambda}_{1997}$ and $\Delta\underline{\lambda}_{08/97} := \underline{\lambda}_{2008} - \underline{\lambda}_{1997}$, respectively; and where the fourth inequality follows when dividing through by $y(2,r)$ the second term in the bracelets, substituting from (8), (9) as well as from (A8), and summarizing.

9 Appendix B

To provide further evidence for the robustness of our results, we estimate the econometric model (11) based on absolute numbers of GPs. Typically, the supply of physicians is measured in per-capita terms. In the presence of interregional migration, however, the use of physicians per capita as a dependent variable may lead to measurement error. At the regional level, it is often observed, that younger people have a greater propensity to migrate. Consider thus a region subject to intense emigration on the part of young individuals. On the one hand, the corresponding decline in population size implies a ceteris paribus decline in the number of physicians per capita. On

ACCEPTED MANUSCRIPT

the other hand, the share of older people increases, generating a positive yet spurious correlation with the number of physicians per capita. Given that the supply of physicians responds to changes in population size only with a lag, a focus on the number of GPs per capita then comprises a measurement error.

To illustrate this point more formally, we decompose the number of GPs per capita (n/ℓ) into a measure of real supply (S) and a measurement error due to interregional migration (ϵ). The regression $\ln(n/\ell) = S + \epsilon = \alpha + \beta X + u + \epsilon$ (with u as an idiosyncratic error term) that should approximate the true model yields a biased parameter β if the explanatory variable X is correlated with ϵ. Since a declining population boosts both the share 60plus and the number of physicians per capita at least in the short run, we can conclude that the parameter corresponding to the share 60plus is likely biased upwards. For this reason we provide here alternative results for the specification (11), using the log of the number of GPs as dependent variable while directly controlling for population size.

A comparison of the results for the number of GPs per capita (Table 3) and the (absolute) number of GPs (Table 6) shows that the bias due to regional population mobility is very small and, indeed, negligible in the most important regressions. Some differences arise when control variables are excluded. Overall, the results are in line with the Hypotheses H1–H6 and our conclusions remain unchanged.

Table 6 about here

10 Appendix C: All Figures and Tables

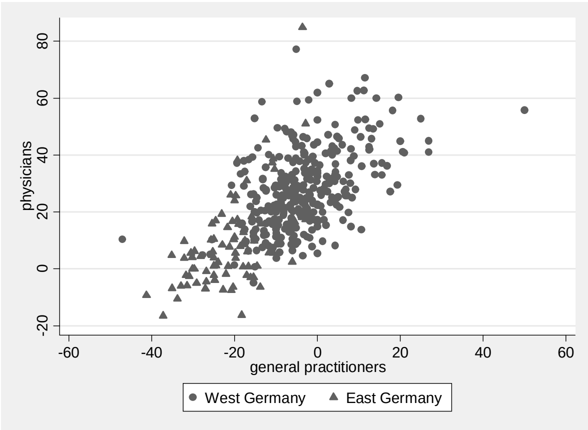

Figure 1: Percentage change, 1995–2009, in the number of physicians and GPs at district level

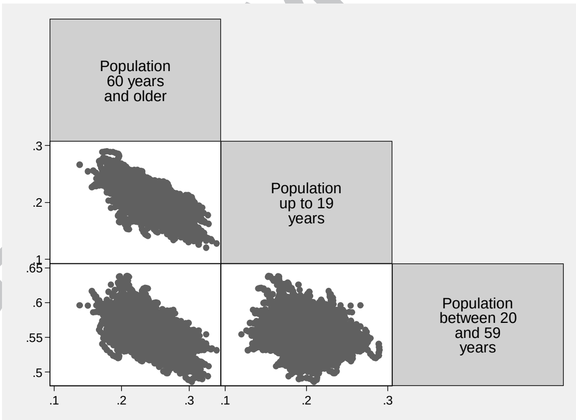

Figure 2: Correlation of population shares at district level

Table 1: Basic Statistics

Variable	Mean	Std. Dev.	Min	Max
GPs per 1000 residents (n/ℓ)	54.027	8.5132	26.1	90
number of GPs (n)	103.89	115.96	9	1918
19minus ($\underline{\lambda}$)	0.2077	0.0251	0.1261	0.2896
60plus ($\overline{\lambda}$)	0.2440	0.0260	0.1608	0.3340
growth of list size ($listgro$)	1.0059	0.0294	0.8385	1.4096
internists per 1000 residents ($intern\_pc$)	19.929	8.8224	0.9	56.5
paediatricians per 1000 residents ($paed\_pc$)	6.9480	2.6215	1.3	19.2
population (ℓ)	200358	226392	34525	3431675
population density ($popdens$)	525.03	670.44	38.1	4270.5
beds for tourism (per 1000 inhabitants)	38.253	51.617	1.5	614.7
share of school leavers with higher education entrance qualification	23.559	8.2058	5.9	67
share of foreigners in population	7.2699	4.6766	0.4	26.3
share of in-commuter amongst all employed	33.872	15.017	7	90.1
GDP per capita (in 1000 €)	24.984	10.038	10.8	86.7
district type I (metropolitan)	0.3335	0.4715	0	1
district type II (municipalized)	0.4354	0.4959	0	1
district type III (rural)	0.2311	0.4216	0	1
excess supply (lower bound)	0.0372	0.1893	0	1
excess supply (upper bound)	0.7565	0.4292	0	1

NOTES: The number of observations is 4890.

Table 2: Basic statistics across district types

	metropolitan districts (type I)	municipalized districts (type II)	rural districts (type III)
GPs per 1000 population	48.167	55.243	60.196
	(7.370)	(6.893)	(7.441)
GPs per km^2	0.416	0.212	0.171
	(0.440)	(0.262)	(0.232)
population density	884.32	378.64	282.25
	(890.4)	(440.5)	(387.6)
60plus ($\overline{\lambda}$)	0.242	0.243	0.249
	(0.022)	(0.028)	(0.027)
19minus ($\underline{\lambda}$)	0.204	0.211	0.207
	(0.020)	(0.027)	(0.027)
growth rate GPs per 1000 pop	-0.041	-0.054	-0.047
	(0.091)	(0.089)	(0.098)
growth rate GPs per km^2	-0.028	-0.065	-0.080
	(0.086)	(0.110)	(0.118)
growth rate population density	0.016	-0.012	-0.035
	(0.056)	(0.063)	(0.071)
growth rate 60plus	0.177	0.175	0.217
	(0.082)	(0.080)	(0.109)
growth rate 19minus	-0.090	-0.133	-0.183
	(0.082)	(0.092)	(0.125)

NOTES: All values are means across districts (and time). Growth rates are calculated between the years 1995 and 2009. Standard errors in parentheses. Type I: 1631 observations; type II: 2129 observations; type III: 1130 observations. For growth rates we have 136 type I districts, 180 type II districts, and 96 type III districts.

Table 3: Main Regressions

Dependent variable: Log of GPs per capita	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
19minus ($\underline{\lambda}$)	-0.452‡	-0.523‡	-0.435‡	-0.488‡	-0.158	-0.220	-0.159	-0.203
	(0.122)	(0.201)	(0.122)	(0.155)	(0.161)	(0.174)	(0.179)	(0.150)
19minus×ruralII		0.215		0.204		0.197		0.180
		(0.139)		(0.154)		(0.145)		(0.162)
19minus×ruralIII		-0.737‡		-0.741‡		-0.683‡		-0.683‡
		(0.223)		(0.215)		(0.218)		(0.211)
60plus ($\overline{\lambda}$)	0.645‡	0.896‡	0.628‡	0.876‡	0.611‡	0.846‡	0.582‡	0.816‡
	(0.060)	(0.046)	(0.056)	(0.042)	(0.050)	(0.050)	(0.049)	(0.044)
60plus×ruralII		-0.342‡		-0.350‡		-0.320‡		-0.334‡
		(0.054)		(0.057)		(0.057)		(0.062)
60plus×ruralIII		-0.999‡		-0.994‡		-0.940‡		-0.933‡
		(0.131)		(0.111)		(0.141)		(0.122)
listgro	0.481‡	0.482‡	0.481‡	0.483‡	0.489‡	0.491‡	0.490‡	0.491‡
	(0.024)	(0.025)	(0.024)	(0.025)	(0.027)	(0.029)	(0.027)	(0.029)
$\ln(intern\_pc)$	-0.063‡	-0.064‡	-0.063‡	-0.064‡				
	(0.015)	(0.015)	(0.015)	(0.015)				
$\ln(paed\_pc)$	0.034‡	0.036‡	0.035‡	0.036‡				
	(0.004)	(0.004)	(0.003)	(0.003)				
$\ln(popdens)$	-0.447‡	-0.482‡	-0.450‡	-0.483‡	-0.421‡	-0.454‡	-0.423‡	-0.454‡
	(0.011)	(0.013)	(0.013)	(0.015)	(0.004)	(0.007)	(0.006)	(0.009)
controls	✓	✓			✓	✓		
within R^2	0.439	0.445	0.438	0.445	0.422	0.427	0.421	0.426
Maddala-Wu test	1999‡	2036‡	3442‡	3889‡	1978‡	2014‡	3656‡	3922‡

NOTES: Significance: ‡=1% and †=5%. Driscoll & Kraay standard errors in parentheses, Maddala-Wu test: unit root test (inverse chi-squared) with the null hypothesis of non-stationarity, specification with dfuller drift lags(1). 4890 obs. Number of regions: 409. Period: 1997–2008. Controls at district level are: GDP per capita, share of foreigners in population, share of school leavers with higher education entrance qualification, tourist accommodation per 100,000 population, and share of in-commuters amongst all employed. All regressions include fixed effects and state-year effects.

Table 4: Regressions accounting for excess supply

Dependent variable: Log of GPs per capita				
	(12)	(13)	(14)	(15)
19minus ($\underline{\lambda}$)	-0.411‡	-0.434†	-0.444‡	-0.563†
	(0.128)	(0.215)	(0.104)	(0.191)
19minus×ruralII		0.226		0.256
		(0.150)		(0.143)
19minus×ruralIII		-0.717‡		-0.597‡
		(0.209)		(0.179)
60plus ($\overline{\lambda}$)	0.679‡	0.883‡	0.475‡	0.688‡
	(0.063)	(0.046)	(0.048)	(0.047)
60plus×ruralII		-0.225‡		-0.294‡
		(0.065)		(0.047)
60plus×ruralIII		-0.860‡		-0.883‡
		(0.118)		(0.072)
$listgro$	0.460‡	0.462‡	0.386‡	0.388‡
	(0.019)	(0.020)	(0.021)	(0.022)
$listgro \times lxs$	0.057‡	0.055‡		
	(0.008)	(0.008)		
$listgro \times uxs$			0.062‡	0.061‡
			(0.005)	(0.006)
$\ln(intern\_pc)$	-0.064‡	-0.065‡	-0.088‡	-0.089‡
	(0.015)	(0.014)	(0.016)	(0.016)
$\ln(paed\_pc)$	0.032‡	0.033‡	0.025‡	0.027‡
	(0.004)	(0.004)	(0.003)	(0.003)
$\ln(popdens)$	-0.418‡	-0.443‡	-0.478‡	-0.512‡
	(0.008)	(0.011)	(0.019)	(0.023)
within R^2	0.455	0.460	0.518	0.523

NOTES: Significance: ‡=1% and †=5%. Driscoll & Kraay standard errors in parentheses, 4890 obs. Number of regions: 409. Period: 1997–2008. Controls at district level are: GDP per capita, share of foreigners in population, share of school leavers with higher education entrance qualification, tourist accommodation per 100,000 population, and share of in-commuters amongst all employed. All regressions include fixed effects and state-year effects.

Table 5: Decomposing changes in GP supply

	metropolitan districts (type I)	municipalized districts (type II)	rural districts (type III)
(1) $\frac{\text{GPs\_pc}_{2008}}{\text{GPs\_pc}_{1997}}$	0.959	0.946	0.953
(2) $\frac{y_{2008}}{y_{1997}}$	1.018	1.018	1.064
of this due to change in			
(3)...age structure	1.041	1.029	1.068
(4)...pop dens	0.992	1.006	1.017
(5)...intern_pc	0.984	0.981	0.977
(6)...paed_pc	1.002	1.002	1.002
(7) $\frac{1+\delta\left(1+g^{\ell}_{2009}\right)}{1+\delta\left(1+g^{\ell}_{1998}\right)}$	1.005	1.009	1.004
(8) $\frac{\overline{v}_{2008}}{\overline{v}_{1997}}$ implied	1.067	1.086	1.121

NOT: Entries in row (1) are taken from the data. Entries in rows (2)–(7) are calculated from equation (12), using relevant coefficients from Table 3 and data from Table 8. Entries in (8) follow from multiplication of the respective entries in rows (2) and (7) and division by the entries in row (1).

Table 6: Regressions based on log number of GPs

Dependent variable: log number of GPs	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)
19minus	-0.445‡	-0.520‡	-0.429‡	-0.485‡	-0.153	-0.218	-0.154	-0.202
	(0.123)	(0.201)	(0.122)	(0.155)	(0.161)	(0.173)	(0.178)	(0.148)
19minus×ruralII		0.221		0.209		0.201		0.185
		(0.140)		(0.155)		(0.145)		(0.162)
19minus×ruralIII		-0.727‡		-0.731‡		-0.672‡		-0.671‡
		(0.222)		(0.214)		(0.217)		(0.209)
60plus	0.638‡	0.886‡	0.620‡	0.865‡	0.608‡	0.840‡	0.578‡	0.808‡
	(0.062)	(0.047)	(0.056)	(0.040)	(0.050)	(0.050)	(0.047)	(0.042)
60plus×ruralII		-0.337‡		-0.345‡		-0.314‡		-0.328‡
		(0.052)		(0.055)		(0.054)		(0.059)
60plus×ruralIII		-0.990‡		-0.984‡		-0.929‡		-0.921‡
		(0.128)		(0.108)		(0.138)		(0.119)
listgro	0.483‡	0.485‡	0.484‡	0.486‡	0.492‡	0.494‡	0.492‡	0.494‡
	(0.023)	(0.024)	(0.023)	(0.024)	(0.026)	(0.028)	(0.026)	(0.028)
ln (*internists*)	-0.063‡	-0.065‡	-0.063‡	-0.065‡				
	(0.015)	(0.015)	(0.015)	(0.015)				
ln (*paediatricians*)	0.033‡	0.035‡	0.033‡	0.035‡				
	(0.004)	(0.004)	(0.003)	(0.003)				
ln (*population*)	0.579‡	0.543‡	0.575‡	0.542‡	0.576‡	0.543‡	0.574‡	0.543‡
	(0.005)	(0.007)	(0.004)	(0.007)	(0.003)	(0.006)	(0.005)	(0.008)
controls	✓	✓			✓	✓		
within R^2	0.574	0.579	0.574	0.579	0.561	0.565	0.561	0.565
Maddala-Wu test	1531‡	1535‡	2867‡	2900‡	1513‡	1520‡	2824‡	2873‡

NOTES: Significance: ‡=1% and †=5%. Driscoll & Kraay standard errors in parentheses. Maddala-Wu test: unit root test (inverse chi-squared) with the null hypothesis of non-stationarity, specification with dfuller drift lags(1). Observations: 4890. Number of districts: 409. Regressions for the period 1997–2008. For control variables see Table 3.

Table 7: Full results for preferred specifications

Dependent variable: Log GPs per capita	(4)		(5)	
	coefficient	S.E.	coefficient	S.E.
19minus ($\underline{\lambda}$)	-0.452‡	(0.122)	-0.523‡	(0.201)
19minus×ruralII			0.215	(0.139)
19minus×ruralIII			-0.737‡	(0.223)
60plus ($\overline{\lambda}$)	0.645‡	(0.060)	0.896‡	(0.046)
60plus×ruralII			-0.342‡	(0.054)
60plus×ruralIII			-0.999‡	(0.131)
listgro	0.481‡	(0.024)	0.482‡	(0.025)
ln (*intern_pc*)	-0.063‡	(0.015)	-0.064‡	(0.015)
ln (*paed_pc*)	0.034‡	(0.004)	0.036‡	(0.004)
ln (*popdens*)	-0.447‡	(0.011)	-0.482‡	(0.013)
beds for tourism (per 1000 inhabitants)	0.012	(0.008)	0.011	(0.008)
share of school leavers with higher education entrance qualification	0.0003	(0.006)	-0.002	(0.006)
share of foreigners in population	-0.004	(0.011)	0.003	(0.012)
share of in-commuter amongst all employed	-0.004	(0.010)	-0.003	(0.011)
GDP per capita (in 1,000 €)	-0.013	(0.015)	-0.015	(0.014)
within R^2	0.439		0.445	

NOTES: Significance: ‡=1% and †=5%. Driscoll & Kraay standard errors in parentheses, 4890 obs. Number of districts: 409. Period: 1997–2008. All controls as logarithms. All regressions include fixed effects and state-year effects.

Table 8: Data for the decomposition analysis

	metropolitan districts (type I)	municipalized districts (type II)	rural districts (type III)
$\overline{\lambda}_{1997}$	0.219	0.220	0.221
$\Delta\overline{\lambda}_{97/08}$	0.0387	0.0385	0.0479
$\underline{\lambda}_{1997}$	0.212	0.223	0.226
$\Delta\underline{\lambda}_{97/08}$	-0.0189	-0.0298	-0.0415
$1 + g^{\ell/n}_{1998}$	0.993	0.988	1.000
$1 + g^{\ell/n}_{2009}$	1.009	1.016	1.013
$\frac{popdens_{2008}}{popdens_{1997}}$	1.016	0.988	0.965
$\frac{intern\_pc_{2008}}{intern\_pc_{1997}}$	1.284	1.352	1.428
$\frac{paed\_pc_{2008}}{paed\_pc_{1997}}$	1.045	1.068	1.044

NOTES: $\Delta\overline{\lambda}_{97/08} := \overline{\lambda}_{2008} - \overline{\lambda}_{1997}$ and $\Delta\underline{\lambda}_{97/08} := \underline{\lambda}_{2008} - \underline{\lambda}_{1997}$.