Price sensitivity and regional variation in health care

Naimi Johansson

School of Public Health and Community Medicine
Institute of Medicine
Sahlgrenska Academy, University of Gothenburg

UNIVERSITY OF GOTHENBURG

Gothenburg 2021

Cover illustration by Ardely

Price sensitivity and regional variation in health care

naimi.johansson@gu.se

ISBN 978-91-8009-200-5 (PRINT)
ISBN 978-91-8009-201-2 (PDF)
http://hdl.handle.net/2077/67121

Printed in Borås, Sweden 2021
Printed by Stema Specialtryck AB

Dedicated to Maj-Lis, Arne, Per and Hjördis

Price sensitivity and regional variation in health care

Naimi Johansson

School of Public Health and Community Medicine, Institute of Medicine
Sahlgrenska Academy, University of Gothenburg
Gothenburg, Sweden

ABSTRACT

Understanding the consequences of current health policy is important in order to design and develop a health care system suitable for future challenges. The purpose of this thesis is to bring evidence on the determinants of regional variation in health care and on individuals' responsiveness to patient out-of-pocket prices in Sweden. The papers included in the thesis are longitudinal register based studies, using representative samples of the Swedish population, with data obtained from national and regional databases. The analyses are primarily based on econometric methods drawing on quasi-experimental approaches to estimate causal effects. The results in Paper I show that regional level mortality and demographics explain a large part of regional variation in visits to specialists, but has limited association with regional variation in visits to primary care physicians. In Paper II, the results show that the relative effect of individual level characteristics outweighs the effect of region-specific characteristics as the drivers of regional variation in pharmaceutical expenditures. The findings in Paper III show that young adults are price sensitive and reduce their use of primary care services after the introduction of patient out-of-pocket prices, with especially strong effects among low-income groups and women. In Paper IV, the findings show that older adults respond to an upcoming elimination of patient out-of-pocket prices by delaying primary care visits in the months before the policy change, but the results show no evidence for a persistent increase in primary care use after the out-of-pocket price elimination.

In conclusion, the findings show that the determinants of regional variation differ within the same health care system, which suggests that the specific institutional settings by type of care are key in understanding regional variation. Further, the results imply that policymakers need to consider heterogeneity and forward-looking behavior in individuals' sensitivity to out-of-pocket prices when developing health care policy.

Keywords: health care utilization, health insurance, regional variation, price sensitivity

ISBN 978-91-8009-200-5 (PRINT)
ISBN 978-91-8009-201-2 (PDF)

SAMMANFATTNING PÅ SVENSKA

Kunskap om effekterna av hälso- och sjukvårdspolicy är viktigt för att bygga och vidareutveckla ett hälso- och sjukvårdssystem för framtida utmaningar. Det övergripande temat för denna avhandling är faktorer som påverkar användning av sjukvård i Sverige med specifikt fokus på att öka kunskaperna om orsaker till regionala variationer i sjukvård och om individers priskänslighet inför patient-avgifter. Regionala variationer syftar till skillnader i användning av sjukvård mellan geografiska områden inom ett land. Priskänslighet i sjukvård handlar om hur individer påverkas av ekonomiska incitament i sjukvårdsförsäkring såsom patientavgifters effekt på användning av sjukvård.

De fyra delarbetena i avhandlingen är longitudinella registerstudier med data hämtad från nationella samt regionala databaser och stickprov baserat på representativa urval av den svenska befolkningen. Analyserna bygger främst på ekonometriska metoder med kvasi-experimentella ansatser i syfte att skatta kausala effekter. Resultaten visar att mortalitet och demografi på regional nivå förklarar en stor del av regionala variationer i besök till specialistläkare, men nämnda faktorer har ett begränsat samband med regionala variationer i besök till primärvårdsläkare. Vidare visar resultaten att regionala variationer i läkemedelskostnader till största del drivs av patienters individuella egenskaper och endast en liten del beror på specifika regionala förhållanden. Gällande priskänslighet visar resultaten att unga vuxna minskar antalet besök i primärvård efter att patientavgifter introduceras vid 20 års ålder, med särskilt starka effekter bland kvinnor och individer från hushåll med lägre inkomster. Resultaten visar också att äldre individer påverkas av en framtida avgiftsfri öppenvård från 85 års ålder genom att minska antalet primärvårdsbesök månaderna innan policy-förändringen, men resultaten uppvisar inga bevis för en permanent ökning i antalet vårdbesök efter att patientavgiften tagits bort.

Sammanfattningsvis tydliggör resultaten från avhandlingen att orsakerna till regionala variationer skiljer sig för olika typer av sjukvård inom ett och samma sjukvårdssystem, vilket tyder på att specifika organisationsstrukturer för respektive typ av vård är viktiga för att förstå regionala variationer. Resultaten från avhandlingen innebär även att beslutsfattare behöver vara medvetna om och ta ställning till att det finns skillnader i hur olika grupper påverkas av patientavgifter, samt att individer är framåtblickande och reagerar även på kommande förändringar i patientavgifter.

LIST OF PAPERS

This thesis is based on the following studies, referred to in text by their Roman numerals.

I. Johansson, N., Jakobsson, N. & Svensson, M. Regional variation in health care utilization in Sweden – the importance of demand-side factors. *BMC Health Services Research*, 2018, 18:403.

II. Johansson, N. & Svensson, M. Regional variation in drug expenditures – evidence from regional migrants in Sweden. Manuscript.

III. Johansson, N., Jakobsson, N. & Svensson, M. Effects of primary care cost-sharing among young adults: varying impact across income groups and gender. *The European Journal of Health Economics*, 2019, 20(8):1271–1280.

IV. Johansson, N., de New, S.C., Kunz, J., Petrie, D. & Svensson, M. Reductions in out-of-pocket prices and forward-looking moral hazard in health care. Manuscript.

CONTENT

ABBREVIATIONS

GDP	Gross domestic product
GLS	Generalized least squares
GRP	Gross regional product
OECD	Organisation for Economic Co-operation and Development
RD	Regression Discontinuity
RKA	Rådet för främjande av kommunal analys (Council for promotion of analysis of local authorities)
SCB	Statistiska Centralbyrån (Statistics Sweden)
SKR	Sveriges Kommuner och Regioner (Swedish association of local authorities and regions)
TLV	Tandvårds- och läkemedelsförmånsverket (Dental and pharmaceutical benefits agency)

1 INTRODUCTION

Health care is something everyone needs from time to time and, as health care most often is financed through common resources in public or private health insurance programs, essentially all members of society contribute financially to the health care system. In order to design and develop the best possible health care given available resources, it is important to understand the consequences of current health policy. One of the main challenges for health care systems today is the high level of expenditures, which has been increasing steadily over the last decades in high- and middle-income countries (OECD 2020). Health care expenditures in 2018 accounted for on average 9% of the gross domestic product (GDP) in high- and middle-income countries, and 11% of GDP in Sweden (OECD 2019). That means that about one tenth of all incomes were spent on health care, and the vast majority of that (84% in Sweden and an average of 71% in high- and middle-income countries) was financed through public funds (OECD 2019). Policymakers need knowledge of how institutional settings, regulations and incentives affect health care utilization and expenditures. The central theme for this thesis is determinants of health care utilization, with specific focus on two topics that have attracted interest in the research literature and are of high policy relevance: regional variation in health care and price sensitivity in health care.

Differences in health care utilization and expenditures across areas within a country, usually referred to as regional variation in health care, have been documented in various health care settings, but it has proven difficult to establish the driving causes of regional variation in health care (Corallo et al. 2014, Cutler et al. 2019, OECD 2014, Skinner 2011). If variations are caused by differences in population health and need for medical care, the variations are not necessarily a problem. If on the other hand, regional variation is driven by unjust allocation or inefficient use of resources, there may be need for improvement (Skinner 2011). In Paper I and II of this thesis, regional variation in physician visits and in pharmaceutical expenditures across the Swedish regions are studied, with aims to determine what factors may explain the variations.

Price sensitivity in health care relates to the way individuals respond to economic incentives in health insurance and to patient out-of-pocket prices. Health insurance lead patients to use more health care then they would if they were to pay the full price of health care themselves (Cutler and Zeckhauser 2000, Pauly 1968, Zweifel and Manning 2000). This is commonly referred to as moral hazard in health insurance, and patient out-of-pocket prices are used as a way to reduce

the scope of moral hazard (Einav and Finkelstein 2018, Pauly 1968, Zweifel and Manning 2000). In Paper III and IV of this thesis, individuals' response to changes in out-of-pocket prices and the impact on the use of primary health care services is studied among young adults and old adults in the Swedish setting. It should be noted that even though the two topics have a common ground, there is no direct (causal) pathway between patient out-of-pocket prices and regional variation in health care. A longitudinal study across Swedish regions found no evidence of a correlation between out-of-pocket prices and the average number of physician visits in the different regions (Jakobsson and Svensson 2016a).

1.1 Theoretical background

1.1.1 Demand for health insurance

The health care market differs from the formalized model of perfect competition even more than the markets for most ordinary goods do, and the main reason is uncertainty. Arrow (1963) described that "all the special features of this industry [the health care market], in fact, stem from the prevalence of uncertainty". There is uncertainty in health and illness, in the sense that an individual cannot determine if, when or how bad she will fall sick and what her need for health care will be. This implies that demand for health care is unpredictable. The risk of illness is also a risk of financial loss, because of high costs of health care and because a reduced ability to make a living often leads to loss of income. In addition to that, there is uncertainty in health care and in recovery from illness, in the sense that the efficacy of a treatment, the quality of the product, is difficult to determine with confidence (Arrow 1963).

Uncertainty and risk in an economic market creates a demand for insurance, and in the case for health and health care there is a demand for health insurance (Arrow 1963, Cutler and Zeckhauser 2000, Pauly 1968). A short note on terminology: from a financial perspective health itself cannot be insured, so the term "health insurance" really refers to insurance for the financial loss of illness (Cutler and Zeckhauser 2000). Even though preventive care such as vaccines can be seen as a real-world applied insurance of health, reducing the risk of disease, but that is really the topic for another thesis.

There is a demand for health insurance because most individuals are risk-averse and prefer an outcome with certainty compared with an uncertain outcome, given the same expected income (Arrow 1963, Cutler and Zeckhauser 2000). The theory is based on the assumptions that an individual's utility is determined by her income, that there is a diminishing marginal utility of income and that the rational

individual seek to maximize her expected utility. From the diminishing marginal utility of income follows that the individual is risk-averse. Thus, when there is a risk of loss of income (due to illness), the individual will have a higher utility of the expected income *I* with certainty under insurance, than the expected utility of (the same) income *I* under uncertainty without insurance. Insurance will lead to a welfare gain to society because spreading (pooling) the risk to a larger population will reduced the total risk (Arrow 1963, Cutler and Zeckhauser 2000, Pauly 1968). Arrow (1963) argued that if the market fails to meet the demand of individuals to insure against the risks of illness, the failure will imply a loss of welfare to society and government intervention will be needed.

1.1.2 Moral hazard in health insurance

The above described theory of demand for health insurance provides an understanding of why health care often is organized in (public or private) health insurance programs. However, even as health insurance results in a welfare gain, it creates other problems as it influences the economic incentives for patients and health care providers, and there is a tradeoff between risk spreading and relevant incentives (Cutler and Zeckhauser 2000). When patients do not pay the full price of health care themselves, moral hazard in health insurance lead patients to demand more health care (Cutler and Zeckhauser 2000, Pauly 1968, Zweifel and Manning 2000). In a broad sense, moral hazard refers to behavioral changes when under insurance coverage, and may in theory take the shape of increased risky behavior and reduced preventive efforts, or increased demand for health services and for new, more costly medical technology (Zweifel and Manning 2000). In the empirical literature, moral hazard in health insurance has come to denote mainly how individuals respond to patient out-of-pocket prices in use of health care services (Einav and Finkelstein 2018). A more general term for consumer responsiveness to price is price sensitivity.

A topic that has gained more interest recently is dynamic incentives and forward-looking behavior in health insurance contracts (Aron-Dine et al. 2015, Einav and Finkelstein 2018, Klein et al. 2020). Many, or perhaps most, health insurance contracts and out-of-pocket schemes vary by the level of expenditures or by age, for example paying the full price out-of-pocket up to a certain level of expenditures or an exemption of out-of-pocket prices up to a certain age. This creates dynamic incentives in the sense that the patient may respond to today's current price or to the future expected price of health care. A rational, forward-looking individual is expected to respond to future price of health care, a behavior which can be refer to as "forward-looking moral hazard" (Aron-Dine et al. 2015, Eliason et al. 2019).

As a measure of the size of price sensitivity it is common to report price elasticities which is calculated as the percentage change in quantity (demanded) given the percentage change in price. Newhouse (2014) have pointed out that the use of elasticities may be misleading in the health insurance context as out-of-pocket prices often are relatively small amounts and with relatively large percentage changes in price, or are considering a change from price zero, which almost by definition will result in a very small elasticity. Instead, Newhouse (2014) recommend to simply describe the responsiveness to out-of-pocket price as the percentage change in quantity.

With regards to price sensitivity, the focus in this thesis is on how changes in out-of-pocket prices impact the use of primary health care services. In Paper III, heterogeneous effects in price sensitivity with respect to sex and income are studied among young adults in the setting of Region Västra Götaland. In Paper IV, the question of forward-looking behavior is raised, considering whether older adults respond in advance to a forthcoming elimination of our-of-pocket prices, in Region Stockholm and Region Västra Götaland.

Definitions

Patient out-of-pocket prices, also known as patient cost sharing, refers to the amount the patient pays directly from her own pocket for health care services, admissions or pharmaceuticals, in contrast to the indirect costs paid by the insurer (the third party payer). Out-of-pocket prices come in many shapes and forms in different health care systems: for example deductibles, copayments and coinsurance rates (Cutler and Zeckhauser 2000). Deductibles (also known as excess) imply that the patient pays the full cost of health care up to a certain deductible limit, where the insurance kicks in, and usually resets on annual basis. Copayment is usually a fixed amount paid for each type of health service. Coinsurance is the term for a percentage rate paid by the patient of the full costs of health care. It is also common with a maximum limit of out-of-pocket spending, often on an annual basis, referred to as stop loss, cap, or out-of-pocket limit.

Price sensitivity (of demand) – consumer responsiveness in demand to changes in price

Moral hazard – in a broad sense individuals' behavioral changes when under insurance coverage, and in the health economics literature mainly in the sense individuals' responsiveness in health care use to out-of-pocket prices

Out-of-pocket prices, cost sharing – general terms for the price paid directly by the patient

Deductibles, copayments, coinsurance – various kinds of out-of-pocket payments

1.1.3 Regional variation

The organization of health care also takes on a perspective of equity and equality. As stated by Cutler and Zeckhauser (2000), health care and health insurance are but means to reach the central goal to promote better health. For example, the goal of the Swedish health care system, according to Swedish law, is good health for the whole population and health care on equal terms (SFS 2017:30). Finding regional variation in health care, where some areas within a country have much higher health care expenditures or utilization compared with other areas, have been seen as a sign of inefficiency in the organization of health care (Skinner 2011). This raises the question of on what grounds regional variation is justified or if all regional variation is unwarranted. The question relates both to the causes and the consequences of regional variation. Empirical evidence from the US have shown that higher health care expenditures did not seem to result in better health outcomes, quality or higher satisfaction (Baicker and Chandra 2004, Fisher et al. 2003, Zhang et al. 2010b).

In this thesis, the focus will be on the driving causes, the determinants, of regional variation. The common way to see the question of what is justified, is that variation caused by differences in health, need for health care and preferences, should not be seen as a problem (Skinner 2011). On the other hand, variation caused by for example differences in allocation of resources, such as more hospitals and physicians located in some areas; a wasteful use of resources, such as high-intensity care based on physician preferences rather than medical need; or

physicians' financial incentives; would be unwarranted regional variation. In a policy perspective, it is relevant to assess how to deal with and reduce unwanted regional variation. If regional variation is primarily driven by place-specific characteristics created by factors like those just described (allocation of resources etc.), policies targeting those factors could reduce regional variation. However, if regional variation is primarily driven by differences in individuals' characteristics, policies with aim to change for example allocation of resources would have little impact on regional variation, or even be counterproductive (Finkelstein et al. 2016). Simplifying, one can say that the individual level characteristics represent typical "demand-side" factors and the place-specific characteristics represent typical "supply-side" factors. Separating the causal effects of "demand" and "supply" have proven very difficult, due to the interdependency between them (Cutler et al. 2019, Finkelstein et al. 2016, Skinner 2011).

Previous evidence, described in more detail in section 1.3, has documented regional variation in health care expenditures, utilization and medical practice within a country, both on an aggregated level (such as total expenditures) and on disease-specific treatment alternatives (Corallo et al. 2014, OECD 2014). Evidence has shown variation across varying geographical units such as regions, provinces, hospital referral regions, and post-code areas. The size of geographical unit matters for describing the size of variations, as a larger number of smaller size units (by definition) implies larger variation (OECD 2014, Zhang et al. 2012). The different measures and the different geographical units of regional variation sometimes makes straight comparisons across studies difficult, but it also shows the importance of understanding regional variation in health care with respect to varying outcome measures and the level of geographical units.

Regional or geographical variation – differences in health care expenditures, utilization or medical practice across geographical areas (such as regions, provinces, hospital referral regions, or post-code areas)

The focus of this thesis is on determinants of regional variation in health care on a structural level, rather than a disease-specific treatment or procedure. Paper I studies what demand-side factors are explaining regional variation in "all cause" physician visits, and Paper II examines whether individual level characteristics or

place-specific characteristics are the main drivers of regional variation in expenditures of prescribed pharmaceuticals. The geographical units assessed are the 21 Swedish regions (corresponding to NUTS3 level by Eurostat standard (Eurostat European Commission 2018)), based on the decentralized organization of health care in Sweden and for reasons of data availability.

1.2 Policy context

Health care in Sweden is organized as a single payer, public health insurance program, funded by taxes and with universal coverage. As already mentioned, stated in Swedish law, the purpose of Swedish health care is to provide good health and health care on equal terms, with priorities based on need (SFS 2017:30). It is a decentralized system where the 21 regions have the responsibility to fund and provide health care services for their residents (Anell et al. 2012). The responsibility for nursing homes and long-term care is assigned to municipal level (290 units).

The last decade and a half, a set of reforms has changed the since 1970's complete public monopoly in health care (Anell 2015). In 2010 the act of free choice reform (SFS 2008:962) increased patient choice and reduced barriers to entry for private providers in primary care. In subsequent years, the reform was expanded to include outpatient specialized care. Currently, both public and private health care providers operate within the publicly funded system, but there are regional discrepancies in the private-public mix. Private health care providers within the publicly financed system and private profits are recurring questions in the public and political debate.

For prescribed pharmaceuticals, decision-making lies on central level where the government authority the Dental and Pharmaceutical Benefits Agency (TLV) determines what medicines will be subsidized. On the pharmacy market, year 2008 marked the start of deregulating the previously state owned pharmacy monopoly, reducing barriers to entry and making over-the-counter pharmaceuticals available outside pharmacies.

1.2.1 Patient out-of-pocket prices

Patient out-of-pocket prices in Swedish health care are relatively low, but with separate policies for outpatient care, inpatient care and prescription pharmaceuticals. To reduce the financial burden for patients who have a higher need of health care there are maximum limits on annual basis. In outpatient care, patient out-of-pocket prices consist of a copayment for each health service provided, and an annual out-of-pocket limit. The copayment amount is set on

regional level and varies depending on level of care (primary or specialized) and health care professional for example physician, nurse or physiotherapist. Figure 1 shows copayments for physician visits in primary and specialized care in each of the Swedish regions in 2020. A majority of regions have set the copayment for a visit to primary care physician to 200 SEK, and for a specialist visit 200–300 SEK (SKR 2020). In Region Västra Götaland, a visit to the primary care physician is 100 SEK and in Region Stockholm 200 SEK. The 12-month rolling out-of-pocket limit for outpatient care is set nationally at 1,150 SEK (in 2012–2018 the cap was 1,100 SEK).

Some groups are excused from out-of-pocket prices: older adults and children (SKR 2020). From age 85 (the 85th birthday), older adults pay no out-of-pocket prices in outpatient care. They still pay out-of-pocket for inpatient care and prescribed pharmaceuticals. The exemption for older adults was implemented nationally in 2017, but some regions such as Region Stockholm preceded the national implementation. There is no national policy on exemption of out-of-pocket prices for children, but most common is that the region offers outpatient care free-of-charge for children and adolescents up to age 20 (the 20th birthday).

For prescription pharmaceuticals, the out-of-pocket scheme takes the form of a 4-step deductible with a 12-month rolling limit of 2,350 SEK (year 2020) set on national level (TLV 2020). In the first step, the patient pays the full price of pharmaceuticals up to 1,175 SEK. Thereafter the patient pays 50% of the costs up to the next level, and so forth in two more steps until the limit is reached.

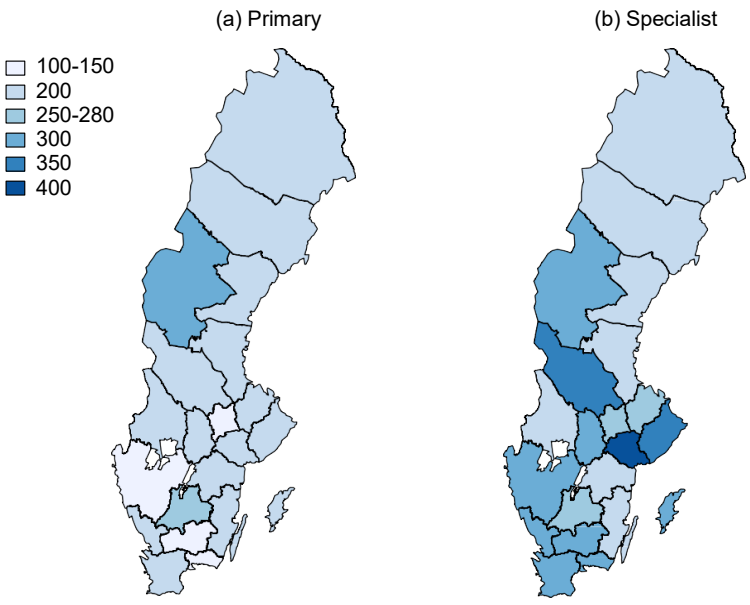

Figure 1. Copayments (SEK) for a physician visit in primary and specialized care
Notes. The copayment amount as of 2020. Maps constructed using data from SKR (2020).

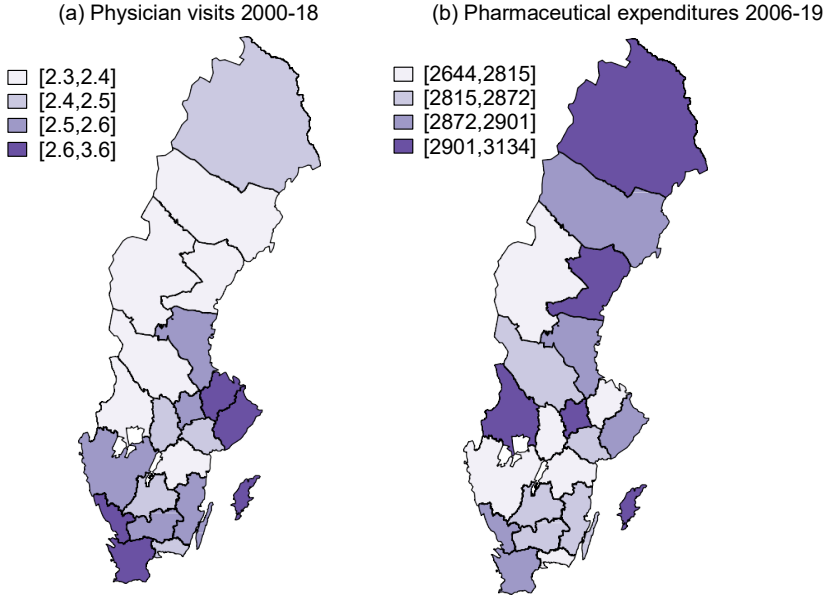

Figure 2. Regional variation in physician visits and in pharmaceutical expenditures
Notes. The averages for each region are pooled over years included. Pharmaceutical expenditures (SEK) refer to costs of prescribed pharmaceuticals bought in pharmacies. Maps constructed using aggregated data available in the online database Kolada (RKA 2020).

1.2.2 How large are the regional variations?

There are notable geographical variations in Swedish health care across the 21 regions. The variations differ depending on outcome measure, for example health care expenditures or number of visits. The maps in Figure 2 show variation across the Swedish regions in the last two decades of a) per capita number of physician visits in outpatient care, and b) per capita expenditures of prescribed pharmaceuticals. Comparing the two maps there is no obvious pattern, it seems the variations in physician visits and pharmaceutical expenditures are unrelated.

Over the years 2000–2018, the average number of physician visits was 2.3 in the region with lowest use and 3.6 in the region with highest use (Figure 2a). The relative difference comparing to the national mean, physician visits ranged from 19% below (Västernorrland) to 28% above (Stockholm) the national per capita number of physician visits (Figure 3). Pharmaceutical spending per capita over the years 2006–2019, ranged from 2,640 to 3,130 SEK (Figure 2b). This corresponds to a relative difference on 7% below (Västra Götaland) to 10% above (Norrbotten) the national mean (Figure 4).

The coefficient of variation, defined as the ratio of the standard deviation to the (unweighted) mean, enables comparison of the size of variations across different outcome units. The coefficient of variation for physician visits was 0.12 and for pharmaceutical spending 0.04, implying that regional variation in physician visits was larger than variation in pharmaceutical spending (Table 1). Values of the coefficient of variation above 0.2, or variation more than two-fold between the lowest and highest using regions are considered high (OECD 2014). Table 1 lists physician visits subcategorized into specialists and primary care physician, showing that variation was larger in specialist visits with a coefficient of variation of 0.17 than in primary care with a coefficient of variation of 0.11. Regional variation in total costs of health care per capita was in line with variations in costs for pharmaceuticals.

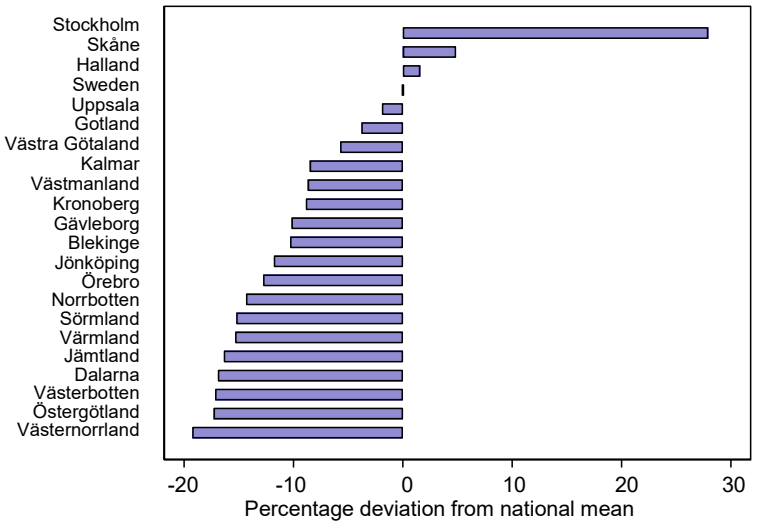

Figure 3. Regional variation in outpatient physician visits: the relative difference

Notes. Zero on the y-axis represent the national (weighted) mean number of physician visits and the horizontal bars show the percentage deviation in mean regional number of physician visits. Data pooled over years 2000–2018. The national mean was 2.8 physician visits per capita per year. Graph constructed using aggregated data available in the online database Kolada (RKA 2020).

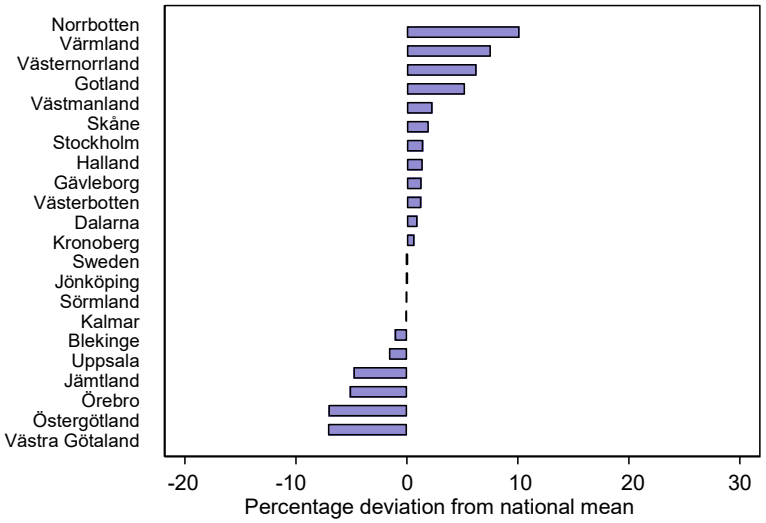

Figure 4. Regional variation in pharmaceutical expenditures: the relative difference

Notes. Zero on the y-axis represent the national (weighted) mean expenditures of prescribed pharmaceuticals per capita and the horizontal bars show the percentage deviation in mean regional pharmaceutical expenditures. Data pooled over the years 2006-2019. The national mean was 2,857 SEK. Graph constructed using aggregated data available in the online database Kolada (RKA 2020).

Table 1. Regional statistics of health care utilization and expenditures, pooled over time

	Physician visits per capita			Health care expenditures per capita (SEK)	
	All	Specialist	Primary	Total	Prescr. pharma.
Data years	2000-2018	2000-2018	2000-2018	2007-2018	2006-2019
Regional					
Mean (Unweighted)	2.56	1.20	1.36	22,774	2,878
St. dev.	0.30	0.21	0.16	1,129	126
Min	2.26	0.97	1.10	21,226	2,654
10^{th} percentile	2.32	1.02	1.20	21,337	2,712
Median	2.51	1.16	1.35	22,813	2,883
90^{th} percentile	2.84	1.50	1.52	23,969	3,039
Max	3.58	1.82	1.76	25,129	3,145
National mean (weighted)	2.80	1.34	1.46	22,630	2,857
Size of regional variations					
Max/min ratio	1.58	1.88	1.59	1.18	1.19
$90^{th}/10^{th}$ ratio	1.23	1.48	1.27	1.12	1.12
Coeff. of var.	0.12	0.17	0.11	0.05	0.04

Notes. The coefficient of variation is defined as the ratio of the standard deviation to the mean. Table based on aggregated data available in the online database Kolada (RKA 2020).

1.3 Previous literature

1.3.1 Regional variation

A large literature covering various scientific fields has documented regional variation in health care. In this (non-conclusive) review, I will provide a background with descriptive evidence of regional variation in health care, shortly touch upon studies of the consequences of regional variation, and then focus on previous literature with aims to explain what determines regional variation.

Descriptive evidence

Already in the 1930's, Glover (1938) noted substantial regional variation in tonsillectomy among schoolchildren in the UK and the US. The starting point of the modern research on regional variation in health care is attributed to Wennberg and Gittelsohn (1973), and their article on small area variations in the state of Vermont, US. The paper includes data of a large set of outcome measures, assigned to 13 hospital service areas. Wennberg and Gittelsohn (1973) showed considerable variation across hospital service areas in resource use such as hospital beds and physicians per capita; in health care utilization such as hospital days and discharges; and in health care expenditures. Supply of physicians was

found concentrated in areas with larger populations, higher per capita incomes and a younger population, which Wennberg and Gittelsohn (1973) marked indicate a poor correspondence between medical need and physician supply.

Since then, the body of literature on regional variation in health care expenditures, utilization and in medical practice has grown steadily. Studies have shown regional variation in productivity in the English NHS, in access to care in France, in mortality and resource use in seven European countries, and in total health care expenditures in Spain (Bojke et al. 2013, Cantarero Prieto and Lago-Penas 2012, Gusmano et al. 2014, Heijink et al. 2015). A multinational report described regional variation in a selected set of health care activities and procedures in 13 countries (OECD 2014). Hospital admissions varied twofold (in some cases even threefold) between areas within Australia, Canada, England, Finland, Italy and Portugal. Within-country variations were highest for cardiac procedures, knee replacement and diagnostic imaging scanning. Cardiac procedures varied more than threefold within Australia, Canada, Finland, France, Italy, Portugal, Spain and Switzerland. The authors of the report concluded that it would be unlikely that such large regional variation was caused solely by differences in morbidity or health (OECD 2014).

In a review of more than 800 studies reporting regional variation in medical practice in high- and middle-income countries, more than half of the studies were from the US and Canada (Corallo et al. 2014). The reviewed studies reported large regional variation for various clinical conditions and surgical procedures, but few of the studies had assessed the causes or consequences of the variations. The majority of research of regional variation in health care is based on data from US Medicare, a public health insurance plan available for people of age 65 years and older. It has been shown that crude rates of US Medicare health service expenditures per beneficiary vary threefold across hospital referral regions (Fisher et al. 2009). The substantial price differences across the US accounted for some of the variation, but was not found to be the main driver of regional variation in Medicare spending (Gottlieb et al. 2010). In data from 2015, where expenditures have been adjusted for differences in price, age, sex and race, a twofold variation remained between the bottom and the top spending hospital referral region (Dartmouth Atlas Project 2020).

Regional variation in unadjusted health care spending seem to be lower in other health care settings compared with the US Medicare. Godøy and Huitfeldt (2020) argued we may expect less regional variation in universal health care systems. For a comparison, in British Columbia, Canada, expenditures in the top-spending region was 50% higher than in the bottom-spending region; the same figure across German counties was 45%; and 24% across the Netherlands' provinces

(Göpffarth et al. 2016, Lavergne et al. 2016, Moura et al. 2019). An exception, with figures closer to those from US Medicare, is Switzerland where unadjusted health care spending in the top-spending canton was 146% of the bottom-spending canton (from numbers in Reich et al. 2012). Some studies have shown that the size of regional variation within a country may vary depending on what type of health care is considered. In Germany, the coefficient of variation was 0.41 for visits to psychotherapist, but 0.12–0.14 for specialist and primary care physician visits (Kopetsch and Schmitz 2014). Zhang et al. (2010a) showed that in US Medicare, the coefficient of variation was 0.08 for pharmaceutical expenditures and 0.12 for (non-drug) medical expenditures, both measures adjusted for age, sex and race. Taken together, these figures suggest that the institutional setting is important for the size of regional variation.

Consequences of regional variation

Part of the literature on regional variation in health care has focused on the consequences of variations. Several studies have shown that higher spending did not produce better health outcomes, quality or higher satisfaction among patients in US Medicare (Baicker and Chandra 2004, Fisher et al. 2003, Zhang et al. 2010b). A potential explanation may be that specialization lead to productivity spillover effects. Chandra and Staiger (2007) found that high-use areas had better returns to certain invasive treatments but reduced returns to alternative treatments, which implied that overall health outcomes were uncorrelated with specialization. Other evidence have shown that higher spending in Medicare did lead to better health outcomes (Doyle Jr et al. 2015). Similarly, Godøy and Huitfeldt (2020) found that Norwegian regions with high hospital spending had modestly better health outcomes compared with low-spending regions. The authors highlighted that policy recommendations aimed to limit regional variation are highly dependent on its impact on health outcomes but that the relationship between regional variation and health outcomes remains unclear with evidence mainly from the US (Godøy and Huitfeldt 2020).

Determinants of regional variation

The driving causes of regional variation in health care is debated, even somewhat of a controversy. Adjusting health care spending for patient characteristics and preferences have been found to explain only a small part (12–18%) of regional variation in US Medicare, which have led authors to conclude that regional level supply-side factors are the main drivers of variations (Anthony et al. 2009, Baker et al. 2014, Sutherland et al. 2009). Specifically physicians' financial incentives, specialization and beliefs about treatment choices have been highlighted as important determinants of regional variation (Birkmeyer et al. 2013, Chandra et al. 2011, Cutler et al. 2019, Fisher et al. 2009, Skinner 2011). Baicker and Chandra (2004) showed that in US states with a high proportion of specialists compared

with general practitioners, costly intensive care of lower quality crowd out effective, low costs quality care. A recent study drawing upon surveys on patient preferences and on physician beliefs about treatment choices, showed that physician beliefs explained more of regional variation in Medicare end of life spending than patient preferences did (Cutler et al. 2019). The proportion of offensive physicians promoting intensive care and defensive physicians encouraging palliative care, were found to explain 36% of regional variation while patient preferences accounted for 20%. Physician beliefs were found to be uncorrelated with organizational and financial incentives (Cutler et al. 2019).

Contrasting the above described focus on supply-side factors, other studies have called attention to the need to fully adjust for demand-side patient characteristics to explain regional variation in US Medicare. Controlling for a wide number of health measures, Zuckerman et al. (2010) showed that health explained 37% of variation between the lowest and the highest quintiles of spending. Adjusting for population health based on diagnoses, 75–85% of variations across areas were explained (Reschovsky et al. 2013). Sheiner (2014) analyzed regional variation using aggregate level data and found that health and socioeconomic factors accounted for most of the variation in spending. Finkelstein et al. (2016) used an innovative empirical approach of regional migration (more below) and concluded that 40-50% of regional variation in Medicare was attributed to individual-level demand-side characteristics.

In Germany, regional variation in health care has been found to be driven mainly by differences in medical need and preferences using a comprehensive morbidity index to account for average health status (Augurzky et al. 2013, Göpffarth et al. 2016, Kopetsch and Schmitz 2014). Variation in hospital utilization across the 16 states could to 56% be explained by differences in health and demographic variables (Augurzky et al. 2013). For different types of physician visits, 29–40% of variation across the 413 counties were explained by health, demography and socioeconomic variables, and up to 70% of state level variation (Kopetsch and Schmitz 2014). Assessing regional variation in total health care expenditures, Göpffarth et al. (2016) found that 55% variation across counties were explained by average health status and demography while factors accounting for medical supply did not add in explaining variations.

In Switzerland, where health care provision is decentralized to the 26 cantons and prices determined within each canton, Schleiniger (2014) showed that regional variation in health care spending is driven by differences in quantity, not in price. Supply-side factors such as density of physicians, proportion of managed care, medical and technological progress, and demand-side socioeconomic factors were found to be significantly related to variation in Swiss health care spending (Reich

et al. 2012). In British Columbia, Canada adjusting health care spending for age, sex, recorded diagnoses and environmental factors, variations were reduced by about 60% (Lavergne et al. 2016).

Regional variation in pharmaceutical expenditures

Pharmaceutical expenditures account for about 20 percent of health care expenditures in high- and middle-income countries (OECD 2019). Despite this, in the literature of regional variation in health care, few studies have documented regional variation in pharmaceutical expenditures or assessed its determinants, but mainly used expenditures (or utilization) of health services as the outcome of interest. Zhang et al. (2010a) showed that pharmaceutical expenditures varied considerably across hospital referral regions in US Medicare even after adjusting for price, demography and health. The authors found that drug expenditures were only weakly correlated to expenditures of (non-drug) health services, which indicates that pharmaceuticals may act both as a substitute or as a complement to health services (Zhang et al. 2010a). In a study of non-prescription pharmaceuticals in Italy, prevalence of disease and per capita income were found to be explaining regional variation (Otto et al. 2018). In studies from various settings, individual characteristics such as age, income and/or education have been found associated with regional variation in antibiotics, painkillers, antidepressants and use of multiple pharmaceuticals (Filippini et al. 2006, Henricson et al. 1998, Hovstadius et al. 2010, Kozyrskyj 2002).

Causal approaches

Most of the evidence on the determinants of regional variation builds on correlation and association; few studies are able to assess the causal pathways. Using the US Medicare eligibility threshold at age 65 and individuals without health insurance pre-65, Callison et al. (2020) estimated the causal effect of supply-side factors in regional variation. The authors found that individuals who gained insurance eligibility in regions with high health care expenditures had a higher increase in health care use than individuals who gained eligibility in low-spending regions. Having adjusted for patient health and demographic measures, authors concluded that the findings were evidence of a causal effect of supply-side factors driving regional variation in US Medicare (Callison et al. 2020).

With aims to tease out the relative effect of place-specific supply-side factors from the effect of individual level demand-side factors as drivers of regional variations, Finkelstein et al. (2016) applied an empirical strategy using patient migration. The method draws on related work using patient migration to decompose the relative effect of health and of physician practice in regional variation in diagnostic records, and physician migration to decompose the relative effect of physician-specific behavior and of environment-specific settings in variation in physician

practice styles (Molitor 2018, Song et al. 2010). Variation in health care utilization in US Medicare was estimated to 50–60% be attributed to a place-specific supply-side effect and the rest 40–50% to an individual level demand-side effect (Finkelstein et al. 2016). The authors showed that the demand-side effect was greater for preventive and emergency care, and lower for diagnostic tests and inpatient care.

A number of recent articles have applied the same empirical approach using patient migration to assess regional variation in fragmented care in US Medicare, in private health care spending in the US, and in physician practice styles in Austria (Agha et al. 2019, Ahammer and Schober 2020, Johnson and Biniek 2020). Using data of the full population of the Netherlands, Moura et al. (2019) showed that about 30% of variation in health care spending across provinces was driven by a supply-side place effect and the rest 70% by a demand-side individual effect. Dividing total health care expenditures by type of health care, the estimated place effect was found to be lower for primary care expenditures but slightly higher for pharmaceutical expenditures (Moura et al. 2019).

Salm and Wübker (2020) estimated that the place effect accounted for about 10% of regional variation in utilization of outpatient services in Germany, and the rest 90% accounted to the individual effect. The authors interpreted the findings as a result of strong restrictions on the supply-side, such as maximum number of physicians by area and deductions for overtreatment; combined with few restrictions in patient choice, such as free choice of physician, no need for referrals, low out-of-pocket prices, low waiting times and low travel time. The place effect was found to be lower for primary care, about 8%, compared with specialist care, about 32%. Extending the decomposition and separating demand-side into observed and unobserved individual level characteristics, 50% of variations were attributed to unobserved factors such as health status and preferences (Salm and Wübker 2020). Regional variation in hospital expenditures in Norway were found to be to 50% attributed to an individual level demand-side effect (Godøy and Huitfeldt 2020). Assessing socioeconomic disparities, Godøy and Huitfeldt (2020) showed that the individual effect was 25% among low educated, 60% among people with upper secondary education and about 100% for people with university degrees.

In summary

Regional variation has been documented in various health care settings in measures of health care expenditures, health care utilization and medical practice. The available evidence suggests that the size of regional variation in health care differ in different health care settings, and possibly depending on what type of health care is considered. One of the overarching aims in the literature has been

to establish the determinants of regional variation, but it has proven difficult to sort out the driving causes. Some researchers conclude that place-specific supply-side factors are the main drivers of variation, while others claim that individual level characteristics on the demand-side play an important role as well. A large part of the literature has focused on various supply-side factors' relation to regional variation, while the impact of specific demand-side factors has been given less attention. Considering the many different measures of outcomes, type of health care studied, levels of geographical units and various methodological approaches, it is not surprising that the evidence is quite mixed and that it is difficult to reach a conclusive consensus. Using patient migration to separate the relative effect of individuals and of place, studies conducted in various health care settings estimate a supply-side place effect ranging from 10–60%. The current evidence suggests that the institutional settings of the health care system plays an important role in understanding regional variation, for both the size of variation and the determinants of variation.

1.3.2 Price sensitivity and moral hazard

There is an extensive literature on price sensitivity and moral hazard in health care, and this (non-conclusive) review will focus on experimental and quasi-experimental evidence with aims to estimate causal effects. I will also describe current evidence on heterogeneity in price sensitivity and forward-looking behavior with respect to patient out-of-pocket prices.

Experimental and quasi-experimental evidence

The evidence from randomized experiments in the field is scarce, but there are two well-known health insurance experiments from the US: the RAND Health Insurance Experiment (HIE) and the Oregon Medicaid experiment. The RAND HIE was conducted in 1974–1981 across six locations in the US, and assigned health insurance plans with various levels of out-of-pocket payments to participating families (Manning et al. 1987). The researchers found a significant response of out-of-pocket prices on health care spending, for example, total expenses were 15% lower in the 25% cost-sharing plan compared with the free plan. Estimated arc elasticities, modelled under a set of assumptions, ranged from –0.14 to –0.43 for various types of medical spending and depending on cost-sharing plans compared (Keeler and Rolph 1988). From Keeler and Rolph (1988) stems the widely cited elasticity of –0.2, in summary of 14 point estimates. Aron-Dine et al. (2013) provided an update of the RAND HIE analysis, reported in contemporary style. After testing threats of validity to the HIE's causal interpretations; non-random assignment to plans, participation and attrition bias, and differential filing of claims; Aron-Dine et al. (2013) could confirm the main

findings of the experiment. However, the authors highlighted that the magnitude of the response of out-of-pocket prices on health care is very uncertain.

The second well-known randomized health insurance experiment was conducted in Oregon, US in 2008–2009 (Finkelstein et al. 2012). In contrast to the RAND HIE, the Oregon Medicaid experiment did not study the response to out-of-pocket prices but assessed the causal effects of health insurance coverage. A lottery was set up where the winners of the lottery won the opportunity to apply for the Medicaid health insurance program, which is aimed towards uninsured low-income individuals. Of the 90,000 people who signed up in the lottery, one third were selected as winners and among them, about 10,000 applied and enrolled in the Medicaid health plan. Finkelstein et al. (2012) showed that health insurance coverage led to increased use of health care services, reduced financial strain and improved self-reported physical and mental health among the treated, compared with the controls who did not gain health insurance coverage. In a two year follow up of a subsample of the original study population, Baicker et al. (2013) found that none of the measured clinical outcomes of physical health differed between the treated and the controls, which implied that insurance coverage had limited effect on health outcomes in the short term.

Policy reforms creating a quasi-experimental setting in the German public statutory health insurance have been used to study the effects of out-of-pocket prices on health care use. Increased out-of-pocket prices for prescription drugs by 50–200% in 1997, were found to reduce demand for physician visits by 10–15% (Winkelmann 2004). At the same point in time, out-of-pocket prices for medical rehabilitation programs were increase by about 100%, reducing demand for these programs by 20–25% (Ziebarth 2010). Estimates of elasticities for rehabilitation programs ranged between –0.3 and –0.5. The evidence from the introduction of out-of-pocket prices for physician visits in 2004 is mixed. A difference-in-differences estimation and a structural model of health care demand with survey panel data showed the out-of-pocket prices had no effect on the number of physician visits (Kunz and Winkelmann 2017, Schreyögg and Grabka 2010). Farbmacher and Winter (2013) on the other hand, using claims data found among young adults a 9% reduction in the number of visits, and an overall decrease in the probability of at least one physician visit by 4 percentage points.

Quasi-experimental evidence from the Netherlands' mandatory social health insurance have shown that the design of out-of-pocket prices matters for the behavioral response (Hayen et al. 2018, Remmerswaal et al. 2019a). Individuals responded stronger to deductibles, which can be seen as a loss, than to no-claim refunds, which can be seen as a foregone gain. A common set back of the empirical design when studying health insurance is the presence of selection

effects, which limits the scope for drawing causal conclusions. Remmerswaal et al. (2019b) took advantage of the Dutch setting where some individuals voluntarily choose a higher deductible, to separate the effect of selection from the moral hazard effect. Individuals who chose a higher deductible had on average lower health care expenditures compared with individuals who paid the standard deductible. The authors found that the difference in spending was a pure selection effect (i.e. of being healthier), and not driven by the higher out-of-pocket costs (Remmerswaal et al. 2019b).

The evidence from a set of quasi-experimental studies from Sweden have shown mixed results. Using a panel data set of the average number of physician visits across Swedish regions, Jakobsson and Svensson (2016a) found no evidence of an impact of the level of out-of-pocket prices. A policy reform in Region Värmland, increased out-of-pocket prices for primary care physician visits from 150 to 200 SEK, was assessed in a difference-in-differences framework with Region Örebro as the control (Jakobsson and Svensson 2016b). With daily level data aggregated from the population, the authors found the policy change had no effect on physician visits in their preferred specification. It might be that the aggregated level data failed to pick up the potential response on individual level. Using detailed individual level data and policy reforms in the age of out-of-pocket price introduction, Nilsson and Paul (2018) found that children and adolescents in Region Skåne, Sweden significantly responded to out-of-pocket payments of 100–300 SEK, increasing the number of physician visits in outpatient care by 5–10% when visits were free of charge.

Heterogeneity in price sensitivity

To get a deeper understanding of price sensitivity in health care, studies have tried to tease out heterogeneity across groups and in types of health care. Most of the evidence is based on subcategorizing the sample into groups by type of health care, health status, income, sex or age; and since for example age and health status are closely correlated, one need to be cautious of when causal interpretation is appropriate.

For heterogeneity by different types of health care, there are mixed results. Increased out-of-pocket prices for children in Taiwan was found to decrease the use of health care services, with largest effect for outpatient visits at teaching hospital, but no effect of inpatient care (Han et al. 2019). The Taiwanese health care system applies free choice of providers without gatekeeping and use differential rates of coinsurance depending on level of specialization of health care. Following a reduction in out-of-pocket prices for older adults in Japan, Shigeoka (2014) and Fukushima et al. (2016) estimated elasticities around –0.2 for both outpatient and inpatient services. When assessing various medical specialties,

treatment types and diagnoses, especially high responsiveness was found in visits for ambulatory care sensitive conditions and for orthopedic and eye specialties (Fukushima et al. 2016, Shigeoka 2014). Low-income groups in the US were found to be more price sensitive with regard to outpatient and emergency services compared with hospital services (Chandra et al. 2014). Studies of price sensitivity of prescription pharmaceuticals have with quasi-experimental approaches estimated elasticities between –0.2 and –0.7 in Denmark and between –0.12 and –0.16 in Quebec, Canada (Contoyannis et al. 2005, Simonsen et al. 2016).

To directly and credibly compare differences in price sensitivity by income groups is very unusual in the literature. Among low-income groups in Massachusetts, US, increased out-of-pocket prices were found to reduce the use of health care services with an overall price elasticity of –0.16 (Chandra et al. 2014). Those results are in line with previous estimates of the general population, but only indirect comparison is possible. Nilsson and Paul (2018) however, have provided evidence from a full population sample of children and adolescents from Skåne, Sweden, and with parental income data on individual level. They showed that the responsiveness to changes in out-of-pocket prices was driven by low-income families, and that the effect among high-income families was close to zero. Similar findings were shown in Dutch data where individuals in low-income areas were found to respond strongly to the introduction of out-of-pocket prices, while individuals in areas of high-incomes did not (Remmerswaal et al. 2019a). Contrasting, Jakobsson and Svensson (2016b) found no discrepancies in price sensitivity for physician visits across different socioeconomic areas, using aggregate level data.

Differential effects with respect to sex have shown mixed evidence, and it is not intuitively straightforward why either men or women would be more price sensitive (at least as long as correlation with an income effect can be ruled out). Evidence from a natural experiment in Norway, where teenagers were excused from an out-of-pocket price of €17.5, showed an increase in the number of visits to primary care physician, 22% increase among girls and 14% among boys (Olsen and Melberg 2018). Similarly, Hayen et al. (2018) found Dutch women responded stronger than men did to the out-of-pocket price. Opposing evidence from Belgium and Germany found men were more price sensitive than women in demand for physician visits (Cockx and Brasseur 2003, Farbmacher and Winter 2013).

Evidence of differential price sensitivity based on health status have often shown that healthier people were more price sensitive and chronically ill people were less price sensitive. This has been found among low-income population in Massachusetts, US, among older adults in Japan and with respect to prescription

pharmaceuticals in Denmark (Chandra et al. 2014, Fukushima et al. 2016, Simonsen et al. 2016). Contrasting, the RAND HIE found no evidence of differential effects of out-of-pocket prices based on health status (Manning et al. 1987). With respect to age, evidence points towards younger people being more price sensitive and older adults being less price sensitive to health care (Farbmacher and Winter 2013, Simonsen et al. 2016). Hayen et al. (2018) however, found no differences between people above age 65 or below (19–64 years) in response to the Dutch cost sharing schemes. There is naturally a strong correlation between age and health status, so it is important to be careful when causal conclusions can be made. A number of papers have studied specifically children or adolescents (mentioned above Han et al. 2019, Nilsson and Paul 2018, Olsen and Melberg 2018), or older adults (more details below), making direct comparison across age groups difficult.

Chandra et al. (2010) showed that increased out-of-pocket payments for prescription pharmaceuticals and physician visits for older adults in the US, lead to a reduction in use of drugs and in visits with price elasticities estimated between –0.1 and –0.2. However, the reductions were offset by increases in hospitalizations. Other studies have used age thresholds in the policy setting to assess the impact of insurance coverage and of changes in out-of-pocket prices among older adults. Card et al. (2008) showed that eligibility to the public health insurance Medicare at age 65 in the US led to increased health care utilization. For low-cost services like physician visits, the increases were largest among groups without health insurance coverage prior to age 65, and high-cost procedures increased primarily in groups that had a supplementary insurance on top of Medicare after 65. In a follow up paper, Card et al. (2009) showed that the eligibility threshold had substantial effects of health outcomes, leading to a reduction in mortality by 20% among emergency patients with particularly acute conditions.

At age 70 in Japan, the coinsurance rate decreases from 30% to 10%, which have been found to lead to increased use of health care services (Fukushima et al. 2016, Shigeoka 2014). In the mandatory health insurance system of Japan, patients have free choice of medical providers and there is no gatekeeping. Neither of the two studies found any effects on short-term health outcomes of the reduced out-of-pocket prices, as measured by mortality, self-reported physical and mental health and by clinical exam outcomes (Fukushima et al. 2016, Shigeoka 2014). In an analysis of how men responded to a reduction in out-of-pocket prices at age 60 in China, hospital admissions were found to increase substantially (Feng et al. 2020).

Forward looking behavior

Whether individuals are forward-looking and respond to dynamic incentives created by thresholds in insurance contracts or health policy reforms has recently gained more interest in the literature recently. Awaiting a forthcoming policy change may create incentives for individuals to respond in advance. Empirical evidence has shown that in many cases, individuals are forward-looking with respect to health care prices and respond in anticipation of future expected price.

At the end of the 1990's in Austria, suspension of the current baby bonus (about €1,100) was announced 10 months prior to implementation, providing incentives for parents to (try) to conceive and give birth before the suspension. Brunner and Kuhn (2014) showed that in the month before the suspension 8% more children were born, but they found no evidence of manipulation of birth dates. The implementation of a more generous insurance contract for older adults in the US Medicare, reduced out-of-pocket costs for prescription pharmaceuticals, was announced two years in advance. Alpert (2016) showed that previous estimates of the implementation effect were overstated, not taking into account the anticipation effect of the forthcoming policy. The announcement itself led to a 6% decrease in use of pharmaceuticals, which suggests a delay in the use of pharmaceuticals in anticipation of the forthcoming policy. The effect was driven by a reduction in the use of pharmaceuticals for chronic diseases but not in the use of pharmaceuticals for acute events (Alpert 2016).

The out-of-pocket price scheme for prescription pharmaceuticals in the US Medicare includes several kink points based on total drug expenditures for the individual. Einav et al. (2015) studied the kink point after which the out-of-pocket price for prescription pharmaceuticals increase, and found evidence of forward-looking behavior and a delay as individuals who were close to the kink point at the end of the year reduced the propensity to claim waiting for the contract to reset at the beginning of next year. Dalton et al. (2020) on the other hand, who studied the same kink point but used a different approach, found evidence of complete myopia – non-forward-looking behavior.

With respect to deductibles in the US, Aron-Dine et al. (2015) used variation in time of insurance enrollment, to show that holding the current price constant individuals responded to the expected end of year price, which was evidence of forward-looking behavior. In a similar manner, Klein et al. (2020) found evidence of forward-looking behavior with respect to deductibles in the mandatory health insurance in the Netherlands, using variations over time in deductible limits. They showed that individuals responded to the expected end of year price, rather than the current price.

Brot-Goldberg et al. (2017) compared individuals who changed from a health plan with no out-of-pocket prices to a high-deductible health plan, and found that consumers responded heavily to the current price, conditional on expected end of year price and prior year's end of year price. Further, the authors found that individuals in the second year responded stronger to the previous year's total expenditures, which suggests individuals learned to be forward-looking in the use of deductibles. Families expecting childbirth during the year are highly likely to reach their family deductible. Guo and Zhang (2019) argued that a forward-looking individual, aware of reaching their deductible, is expected to have a smooth spending over the year. However, they found that fathers in families expecting childbirth increased medical spending by 11% per month once the deductible was reached, rejecting the null of full forward-looking behavior.

In summary

The general understanding in the literature of moral hazard and price sensitivity in health care is that price of health care matters. The available experimental and quasi-experimental evidence has proved that health insurance coverage and out-of-pocket prices causally affect the level of health care utilization: Under insurance coverage people use more health care services than without coverage and with lower out-of-pocket price people increase the use of health care services. The magnitude of price sensitivity is usually said to be small, as estimated price elasticities have been found to be inelastic between –1 and 0, but the size of the effect is very uncertain.

The current evidence suggests that there is heterogeneity and forward-looking behavior in the responsiveness to health care prices, but there is more to learn about what mechanisms are driving such effects. Some of the German and Swedish evidence show no responsiveness to changes in out-of-pocket prices in full coverage health care systems. One may note that this is not evidence against the conclusion stated above ("price of health care matters"). Rather the evidence suggests, that in a health care setting where services are provided almost free-of-charge, the average effect of low out-of-pocket prices is limited. While as other evidence shows, some groups of people will respond even to small changes in out-of-pocket prices, for example younger and healthier individuals or individuals with low incomes bound by tight liquidity constraints.

1.4 Rational for the thesis

While there are many interesting and relevant issues to raise related to health policy and the determinants of health care utilization, the specific topics of this thesis are price sensitivity and regional variation in health care. In a Swedish

perspective, the studies included in the thesis are important and policy relevant on their own since there is a lack of scientific evidence of determinants of regional variation and of the impact of out-of-pocket prices in the Swedish health care setting. In an international perspective, a lot of scientific evidence especially on regional variation in health care but also on price sensitivity are based on studies from the US. Providing evidence from various types of health care settings is valuable and important to gain a deeper understanding with respect to both these topics.

In the previous literature on regional variation in health care, the evidence from a single payer, universal coverage national health care system is limited, to my knowledge, to one study (Godøy and Huitfeldt 2020). There has been less focus on the importance of demand-side factors compared with the supply side, perhaps due to lack of data availability. The literature has given little attention to differences by type of care, and whether regional variation in for example outpatient care, inpatient care and pharmaceutical expenditures are driven by the same factors. The evidence is specifically scare on the drivers of regional variation in pharmaceutical expenditures.

In the preceding literature on price sensitivity and moral hazard in health care, evidence is still scarce on how different socioeconomic groups respond to out-of-pocket prices, especially with respect to income, where selection effects and lack of data often make such analyses difficult. The evidence of how older adults respond to out-of-pocket prices is restricted to a few settings, despite them being high consumers of health care and thus a policy relevant group to gain more knowledge about. There is in general limited causal evidence from full coverage health care systems with low to moderate out-of-pocket prices. Previous evidence of forward-looking behavior in health care is restricted to the responsiveness to dynamic incentives created by deductibles, spending kinks and policy reforms, while there is lack of research on how age thresholds in out-of-pocket schemes impact health care utilization. Few studies have assessed whether a forthcoming more generous policy lead individuals to delay health care in advance, and what consequences such a delay could have.

This thesis contributes to the scientific literature by the following. On a general note, access to rich individual level register data has advantages that are rare in an international perspective. Register data of a representative sample of the full population enables to make direct comparison for example across socioeconomic groups. Further, the high level of detail in the data, such as income on individual level and specific dates of birth and of visits, provides the opportunity to estimate effects with very high precision.

Paper I and II in the thesis specifically contribute with evidence of to what extent demand-side factors explain regional variation in outpatient physician care and whether individual level characteristics or region-specific characteristics are the main drivers of regional variation in pharmaceutical expenditures. The hypotheses are that mortality, demographic and socioeconomic variables explain the major part of regional variation in physician visits, and that individual level demand-side factors are the main drivers of regional variation in pharmaceutical expenditures in Sweden.

Paper III and IV in the thesis specifically contribute with estimates of heterogeneous effects in price sensitivity with respect to income and sex, estimates of price sensitivity among older adults (85-years-old), and quantifying the effects of forward-looking behavior with respect to out-of-pocket prices in primary care. The hypotheses are that young adults reduce the number of physician visits when they need to pay the out-of-pocket price as compared to health care free-of-charge, that older adults delay primary care visits awaiting the policy for free-of-charge health care and that older adults increase the number of primary care visits when health care is free-of-charge. Paper IV additionally contributes to the literature by developing a methodological framework to assess the presence of health care delays awaiting a more generous out-of-pocket price policy.

2 AIM

The overall aim of the thesis is to bring evidence on the determinants of regional variation in health care and on individuals' responsiveness to patient out-of-pocket prices in Sweden.

The specific aim for each study is

Paper I To study the importance of demand-side factors in explaining regional variation in outpatient physician visits

Paper II To determine the relative effect of individual level demand-side factors and of region-specific supply-side factors as drivers of regional variation in pharmaceutical expenditures

Paper III To estimate the effects of introducing patient out-of-pocket payments on primary care use, and to study heterogeneity based on sex and income

Paper IV To assess the impact of eliminating patient out-of-pocket payments on primary care use, and to quantify the effects of forward-looking behavior with regards to future price of health care

3 DATA

3.1 Sample and data sources

All four papers in the thesis are longitudinal register based studies. Each of the studies use a different data set and sample, but all are based on the Swedish population and have been obtained from Swedish national and regional databases. The databases contain long time series of individual level data and can be linked to each other or other registers thanks to the unique identifiers used by Swedish authorities. Table 2 gives a summary of the sample, years and data sources for each study.

In Paper I, with focus on regional variation in physician visits and the association with typical demand-side factors, regional level data of health care, demography and socioeconomic variables were collected from publicly available online sources of the Swedish municipal and regional database (Kolada), the National Board of Health and Welfare, Swedish Association of Local Authorities and Regions, and Statistics Sweden. Data of the 21 Swedish regions (corresponding to NUTS3 level by Eurostat standard (Eurostat European Commission 2018)) over years 2001–2014 were used. All the regional data included are aggregates of full population administrative records. Assessing regional variation in pharmaceutical expenditures in Paper II, individual level data from the register of prescribed pharmaceuticals of the National Board of Health and Welfare were used. The data set consisted of a random sample of 1,000,000 Swedish inhabitants and the data covered years 2007–2016.

The National Board of Health and Welfare holds registers, besides prescribed pharmaceuticals, of health care utilization such as hospital admissions and outpatient visits to specialists. For the two papers with focus on price sensitivity and assessing changes in out-of-pocket prices for outpatient care, it was important to include both primary and specialized outpatient care. There is however, no national register on primary care utilization, but some regions keep detailed individual level databases of health care use, including primary care. Thus, data were obtained from Region Västra Götaland's *Vega register* (Paper III and IV) and from Region Stockholm's *VAL database* (Paper IV). In Paper III, the full population of 18–21 year-olds in Region Västra Götaland over the years 2014–2015 made up the sample (73,000 individuals). In Paper IV, the full population of 81–87 year-olds in Region Stockholm and in Region Västra Götaland over the years 2014–2018 made up the sample (40,000 individuals).

For the three individual level studies (Paper II, III and IV), the attainted health care data were linked with demographic and socioeconomic background data from Statistics Sweden's registers. The detailed records from Statistics Sweden, for example being able to include education and income on individual level in the analyses, are an advantage in international comparison where for example income data usually are based on surveys or on aggregated numbers.

Table 2. Summary of data used in each paper

Paper	Sample	Years	Sources
I	Regional level aggregated data based on the full population N=21 regions	2001-2014	Swedish municipal and regional database (Kolada), the National Board of Health and Welfare, Swedish Association of Local Authorities and Regions, and Statistics Sweden
II	Random sample of the Swedish adult population N=1,000,000 individuals	2007-2016	Register of prescribed pharmaceuticals from the National Board of Health and Welfare linked with background variables Statistics Sweden
III	All residents in Region Västra Götaland, born 1993-1996 N=73,000 individuals	2014-2015	Vega register from Region Västra Götaland linked with background variables Statistics Sweden
IV	All residents in Region Stockholm and Region Västra Götaland, born 1931-1933 N=40,000 individuals	2014-2018	Vega register from Region Västra Götaland and VAL database from Region Stockholm linked with background variables from Statistics Sweden

3.2 Variables in use

In Paper I, regional variations of two outcome variables were assessed: number of visits to primary care physician and number of visits to specialist, defined as the regional annual average. The independent variables included were mortality rate; a set of variables for demography (proportion of women and of older adults); a set of variables measuring social capital and economic structure (proportion of educational attainment, gross regional product per capita, proportion of unemployment and average level of financial assistance); and a set of variables of health care resources (number of physicians per capita, number of primary care centers per capita, and public-private mix of primary care; in Paper I referred to as supply side variables).

In Paper II, the outcome variable was expenditures of prescribed pharmaceuticals per year, defined as the sum of the cost for the payer (the region) and the patient's out-of-pocket cost. The most important variable for the empirical identification strategy was the region of residence for each year to be able to categorize individuals who moved, and when and where they moved. Age, sex, individual level income, marital status and the number of children in household were included as independent variables.

For the two papers on price sensitivity, Paper III and IV, the outcome of interest was visit in outpatient care, with specific focus on primary care. For comparison other categories of health care use such as visits to specialist, visits to non-physician health care professionals and hospital admissions were assessed. For the particular empirical method used, the Regression Discontinuity design, it was an advantage to have detailed data of date of visit and the age at point of visits. In Paper III, week of birth was used to calculate the individuals' age at visit in weeks, and in Paper IV, the exact date of birth was used enabling to calculate age at visit in days. Sex and income, defined as equivalized household income (household income adjusted for size and age of household members), were used to run subgroup analyses in Paper III.

3.3 Ethical considerations

Working with individual level register data comes with responsibility. The data sets contain specific information about thousands of individuals and it is of great importance to use the data only for the specified research purpose, and to protect the data from unauthorized access. For all individual level data used in the thesis, the personal identification numbers were replaced by anonymous observation numbers before the data were delivered to the research group. The regional ethics review board in Gothenburg approved the merging of registers and the analysis plans for Paper II (#803-17), for Paper III (#359-16) and Paper IV (#185-18). No ethical review was necessary for Paper I, as the study did not include any individual research subjects.

4 METHODS

When evaluating policy effects, one needs to be careful not to mistake correlation for causality. In search for the true causal effect of a policy or treatment, the main problem is that one can never actually know the counterfactual. The gold standard for drawing inference of the causal treatment effect is randomized controlled trials, common in clinical and laboratory research but it is difficult and often unethical to conduct large scale randomized experiments in social, economic and health policy (Athey and Imbens 2017). Thus, in policy evaluation one often need to rely on observational data and search for sources of exogenous variation other than from controlled random assignment, such as natural experiments or quasi-experimental settings. Applied econometrics provide empirical methods to tease out causal effects from such settings, but it is important to be aware of the methods' assumptions and limitations when interpreting results.

In this chapter, the empirical methods used in the thesis are described. In Paper I, the setting provides no source of exogenous variation and the analysis, performed in a random effects model, is based on associations. In Papers II, III and IV econometric models with potential to draw causal conclusions are used. In Paper II, fixed effects models combined with exogenous variation from regional migration is used in a decomposition analysis and an event study analysis. In Paper III and IV, regression discontinuity (RD) design forms the base of the analyses, with an extension of RD in combination with a donut regression and a regression kink design.

4.1 Random effects

Assessing to what degree potential determinants can explain regional variations in physician visits in Paper I, a random effects model estimated by generalized least squares (GLS) is applied. The regression equation is specified as

$$y_{it} = \alpha + \beta X_{it} + \delta_i + \varepsilon_{it} \qquad (1)$$

where y_{it} is physician visits per capita in region i in year t, α is a constant, X_{it} is a vector of explanatory variables, β is a vector of coefficients representing the marginal effect of each covariate, δ_i is a random effect with region-specific intercepts and ε_{it} is an error term (Wooldridge 2014). The random effects and the error term are assumed to be normally distributed with mean zero and variance σ_δ^2 and σ_ε^2, respectively (Bell and Jones 2015). The GLS estimates the variances of

the two disturbance terms: $\hat{\sigma}_{\delta}^{2}$ captures the time-invariant variation between regions and $\hat{\sigma}_{\varepsilon}^{2}$ captures the remaining variation over time (within regions). Successively adding sets of independent variables; mortality, demographic, socioeconomic and health care resource variables; the amount of regional variation explained is assessed by the reduction in the estimated standard deviation of the random effects, $\hat{\sigma}_{\delta}$, in each step.

4.2 Fixed effects and regional migrants

To tease out whether individual level "demand-side" factors or place-specific "supply-side" factors are the main drivers of regional variations in pharmaceutical expenditures in Paper II, an econometric approach using individuals who move across regional borders is used (Finkelstein et al. 2016). Regional migration is assumed to create an exogenous variation of health care setting for the individual (the exogeneity assumption will be violated if the individual moves because of health care need). To explain the intuition behind the approach, consider an individual who moves from an on average high use region to an on average low use region. If the migrant uses health care on the same high level after the move as before, place-specific supply-side factors would seem to be irrelevant in determining the migrant's use of health care, and individual level demand-side factors explaining the level of health care use. If, on the other hand, the migrant adjusts and reduces her level of health care use in the new low-use region, regional level supply-side factor are important drivers of regional variations.

Based on regional migration, two different types of fixed effects models are applied to estimate how large share of variation can be explained by place-specific supply-side factors, and how much is driven by individual level demand-side factors (Finkelstein et al. 2016). Recent studies have built on this approach to assess regional variation in various type of health care outcomes and in different settings (Godøy and Huitfeldt 2020, Moura et al. 2019, Salm and Wübker 2020).

4.2.1 Decomposition analysis

In a three-way fixed effects model with region, individual and year fixed effects, the estimated region fixed effects are used to decompose regional variation into one part attributed to region and one part attributed to individuals. The regional migrants are the key component in the identification, because if each person lived in the same region throughout, individual fixed effects and region fixed effects would be perfectly correlated. The regression equation is specified as

$$y_{ijt} = X_{it}\beta + I_R + \rho_j + \tau_t + \alpha_i + \varepsilon_{ijt} \qquad (2)$$

where y_{ijt} is log expenditures of prescribed pharmaceuticals of individual *i* in region *j* in year *t*. X_{it} is a vector of individual characteristics with parameter vector β, and I_R is a vector of binary indicators for the number of years since migration. The three levels of fixed effects are regional ρ_j, time τ_t and individual α_i. The error term ε_{ijt} is time-varying individual level disturbance (Finkelstein et al. 2016). To decompose regional variation into a region effect and an individual effect, the estimated region fixed effects $\hat{\rho}_j$ and the regional average expenditures $\bar{y}_j$ are used. Comparing regions A and B (or two groups of regions), the region effect of the difference in average expenditures between A and B is calculated as the decomposed share

$$S_{region} = \frac{\bar{\hat{\rho}}_A - \bar{\hat{\rho}}_B}{\bar{y}_A - \bar{y}_B} \qquad (3)$$

The individual effect's share of regional variations is then $S_{ind} = 1 - S_{region}$. Confidence intervals are estimated by bootstrapping the sample.

4.2.2 Event study

A different way to separate regional variation into region effect and individual effect, is a kind of event study analysis (Finkelstein et al. 2016). The event study make use of regional migrants only, and assess to what extent average regional expenditure determine individual expenditures, when moving from one region to another. To do this, the variable D_i is defined as the difference in average log expenditures between the region of origin and the region of destination $D_i = \bar{y}_{j_{dest}(i)} - \bar{y}_{j_{orig}(i)}$ for individual *i* who moves from region j_{orig} to region j_{dest}. D_i is included as the main independent variable in a two-way fixed effects model with time and individual fixed effects. The regression equation is specified as

$$y_{it} = D_i I_{t>r}\theta + X_{it}\beta + I_R + \tau_t + \alpha_i + \varepsilon_{it} \qquad (4)$$

where y_{it} is log expenditures of prescribed pharmaceuticals for individual *i* in year *t*, $I_{t>r}$ is a binary indicator for before (0) or after (1) the move, θ is the main parameter of interest and the other variables and parameters are defined as described above. The parameter θ represents the change in expenditures at the time of migration, given the (approximate percentage) difference in expenditures between origin and destination region (D_i). θ is interpreted as the share of regional variation attributed to a place-specific region effect (Salm and Wübker 2020). If $\theta = 0$, the difference in expenditures between regions does not affect individual expenditures, and if $\theta = 1$ the difference in regional average expenditures completely predicts individual expenditures.

4.3 Regression discontinuity design

In the two papers on price sensitivity, the age threshold policies in the out-of-pocket scheme provides quasi-experimental settings to estimate the causal effect of the price change in a RD design. In Paper III, the introduction of out-of-pocket prices (copayments) at the 20th birthday is assessed in a standard RD design, and in Paper IV, the elimination of out-of-pocket prices at the 85th birthday is assessed in a development of the RD in combination with a regression kink design and a donut regression.

The key identification assumption of the RD design is that the outcome, visits in outpatient care, is a smooth function a running variable, age, which in turn discontinuously determines the treatment, to pay or not pay the out-of-pocket price (Angrist and Pischke 2014, Lee and Lemieux 2010). Intuitively, individuals just below and individuals just above the age threshold, are assumed to have very similar health care needs, and the only thing that differs between them is the out-of-pocket price. In another perspective, working with longitudinal data sets, health care need of an individual is assumed to be very similar in days just before their birthday and just after. If no other determining factors change discretely at the threshold, and there is no manipulation, the discontinuous change in health care use at the birthday threshold can be interpreted as the causal effect of the change in out-of-pocket price (Angrist and Pischke 2014, Lee and Lemieux 2010). To formalize, following the common practice in the literature (see e.g. Card et al. 2009, Lemieux and Milligan 2008), consider the regression equation

$$y_{it} = f(Age_{it}, \delta) + \beta_1 PostBd_{it} + \varepsilon_{it} \qquad (5)$$

where y_{it} is the number of outpatient visits for individual *i* in time period *t*. The time periods are defined as weeks (Paper III) or days (Paper IV) centered around the birthday threshold. Narrow time measurement is an advantage in the RD framework, as it allows to get very close to the threshold. $f(*)$ is a smooth function of the running variable Age_{it}, centered around the individual's birthday, and of the parameter vector δ. $PostBd_{it}$ is a binary indicator taking the value 0 before the birthday threshold and the value 1 after. In varying specifications, Age_{it} may take a simple linear form or be extended to a second and third degree polynomial. To allow the slope of the running variable to differ before and after the threshold, an interaction between the $PostBd$-indicator and the Age-polynomial is included. The parameter of interest is β_1, estimating the discontinuous change at the threshold i.e. the causal effect of the price change. ε_{it} is an error term.

4.3.1 Donut regression

As mentioned, one of the underlying assumptions for the causal interpretation of the RD model is that there is no manipulation around the threshold (Lee and Lemieux 2010). The assumption will be violated if individuals are forward-looking and adjust their health care seeking behavior in advance of the policy change. With this type of manipulation around the threshold, the estimated jump at the threshold will represent a combination of the "pure", long-term price effect and an anticipation effect of forward-looking behavior. In Paper III and in several similar RD-applications, a so called "donut RD" is used to rule out potential anticipation effects of the future policy (Barreca et al. 2011, Fukushima et al. 2016). The practice simply excludes observations close to the threshold creating a donut hole, a gap, around the policy threshold and estimates the "pure" treatment effect by eliminating the possibility of short-term adaptions close to the threshold.

4.4 Donut RD with kink

In Paper IV, the presence of forward-looking moral hazard and a potential delay effect is studied by developing the donut RD design in combination with a regression kink design. The regression kink design is similar to the RD framework, but assess the change in slope, a kink, at a certain threshold (see Card et al. 2015, Simonsen et al. 2016 on regression kink).

Facing an elimination of out-of-pocket prices at the 85th birthday: if individuals are forward-looking they have incentives to delay health care use until after the birthday threshold has passed. Approaching the birthday threshold, at a certain point in time, the benefit of delaying a health care visit will exceed the cost of delaying (i.e. save the out-of-pocket payment but perhaps have to endure pain or risk of worsened health). At this time point, the number of visits will start to decrease and create a break, a kink, in the underlying trend in visits as individuals enter the donut hole of delay. At the birthday threshold, the immediate price drop is expected to result in a discontinuous jump of increased visits, both additional visits due to the lower price and the delayed visits shifted until after the threshold. Some time after the threshold, the visits that were delayed and shifted until after the birthday will be fewer, the number of visits will decrease until a new post-policy trend is stabilized. This will generate a second kink, leaving the donut hole.

The donut RD kink model is formalized in the following regression equation, flexible to allow for a discontinuity at the threshold and additionally kinks pre and post policy

$$y_{it} = f(Age_{it}, \delta) + \beta_1 PostBd_{it} + \beta_2 Period2_{it}(Age + m)_{it} + \beta_3 Period3_{it}(Age - n)_{it} + \varepsilon_{it} \quad (6)$$

where $Period2_{it}$ and $Period3_{it}$ are binary indicator variables representing the donut hole pre-policy period 2 and the donut hole post-policy period 3. $(Age + m)_{it}$ and $(Age - n)_{it}$ are complementary running variables of Age starting m and n days away from the birthday threshold, depending on the size of the donut. The other variables and parameters are defined as above. The main parameters of interest are β_2 and β_3, estimating the kinks in the underlying trend pre and post policy. β_1 estimates the "pure" price effect as the persistent change in outpatient visits due to the out-of-pocket price elimination.

5 RESULTS

5.1 Explaining regional variation in physician visits

In Paper I, estimating how much regional variation in physician visits in Sweden is explained by four sets of independent variables; the results show that regional variation in visits to primary care physician has little association with regional mortality and demographic factors as these two sets of variables marginally *increase* the estimated regional level variation, $\hat{\sigma}_\delta$ (Table 3). Successively adding a set of socio-economic variables explains 11% of regional variation and a total of 33% of variation is explained when adding a set of health care resource variables. For variation in visits to specialists, the results are almost the reverse. Mortality and demographic variables together explain about 50% of regional variation in specialist visits, while successively including socioeconomic and health care resource variables only marginally increase the degree of explanation. In total, the included variables explain a larger share of variation in specialist visits (50%) than in primary physician visits (33%). A large part (50–67%) of regional variation in physician visits remain unexplained by included variables.

Table 3. Estimated regional variation and degree of explanation in physician visits

	Visits to primary physicians		**Visits to specialists**	
	$\hat{\sigma}_\delta$	% of $\hat{\sigma}_\delta$ explained	$\hat{\sigma}_\delta$	% of $\hat{\sigma}_\delta$ explained
Unadjusted	0.1597	..	0.2152	..
Adjusted for				
Mortality	0.1580	1.1%	0.1761	18.2%
+ Demography	0.1652	–3.4%	0.1086	49.5%
+ Socio-economy	0.1418	11.2%	0.1025	52.4%
+ Health care resources	0.1064	33.4%	0.1069	50.3%

Notes. $\hat{\sigma}_\delta$ is the estimated standard deviation of random effect δ_i (variations on regional level). % of $\hat{\sigma}_\delta$ is the percentage of regional variations explained by included covariates. Years included in analyses: for primary care 2002-2014 and for specialists 2001-2013. Adapted from Paper I (Johansson et al. 2018).

5.2 The drivers of regional variation in pharmaceutical expenditures

Assessing the drivers of regional variation in pharmaceutical expenditures in Sweden in Paper II, the decomposition analysis estimates the share of regional variations attributed to a region effect to 0.09 with confidence interval –0.06 to 0.23, when comparing regions above and below the median in the preferred model specification (Table 4, column 2). The share of variation driven by an individual effect is estimated to 1 – 0.09 = 0.91 and confidence interval 0.77 to 1.06. The confidence interval for the region share overlaps zero, which means that we cannot reject the null hypothesis of no region effect. With an upper limit of 0.23, we interpret these results as the region share may be small but is certainly less important than the individual share, and thus individual level demand-side factors are the main drivers of regional variations in pharmaceutical expenditures.

The results of the decomposition analysis are dependent on what groups of regions are being compared. When comparing the regions in the top and bottom spending quartiles the estimated region effect is 0.05, somewhat smaller than in the main comparison, and the individual effect is slightly larger. This implies that the individual effect is an even more important driver of regional variation when comparing the highest spending quartile of regions to the lowest spending quartile.

The results in the event study analysis are similar, with an region effect ($\hat{\theta}$) of 0.09 and a confidence interval of –0.08 to 0.27 in the preferred model (Table 5, column 2). A region effect of zero would suggest that the differences in average regional expenditures has little effect on individual expenditures after moving to a new region. The confidence interval upper limit of 0.27 implies that even if there is a positive region effect, the individual level effect ($1 - \hat{\theta}$) clearly outweighs the region effect. Taken together our results show that individual level characteristics are the main determinants of regional variation in pharmaceutical expenditures in Sweden.

Table 4. Regional variation of pharmaceutical expenditures – Decomposition analysis

	Model 1	Model 2
Comparing regions above and below the median		
Difference in average log expenditures	0.148	0.148
Region share (95% CI)	0.064 (–0.082; 0.215)	0.085 (–0.063; 0.231)
Individual share (95% CI)	0.936 (0.785; 1.082)	0.915 (0.769;1.063)
Comparing top and bottom quartile of regions		
Difference in average log expenditures	0.221	0.221
Region share (95% CI)	0.066 (–0.093; 0.220)	0.054 (–0.096; 0.214)
Individual share (95% CI)	0.934 (0.780; 1.093)	0.946 (0.786; 1.096)
Independent variables	No	Yes
Year FE, region FE, years since move	Yes	Yes
No of ind-year obs.	7,830,395	7,830,395
No of ind.	910,639	910,639

Notes. The effect shares are estimated in fixed effects regressions with ln(exp+1) as the dependent variable and using the 95-trimmed sample of both migrants and non-migrants. For each model, the decomposition is estimated comparing regions above/below median expenditures, and the top and bottom quartile of regions. Confidence intervals are estimated by bootstrapping with 250 repetitions drawn at the individual level (CI:s constructed by 2.5 and 97.5 percentile of the bootstrap estimates). In Model 1 the regression is run without independent variables of individual characteristics. In Model 2, independent variables include indicators for age-gender group, individual income, marital status and number of children in the household.

Table 5. Regional variations of pharmaceutical expenditures - Event study

	Model 1	Model 2
$\widehat{\theta}$ (st.err.)	0.089 (0.088)	0.094 (0.088)
95% C.I.	–0.083; 0.261	–0.077; 0.266
Independent variables	No	Yes
Years since move	Yes	Yes
Year FE	Yes	Yes
No of ind-year obs.	491,378	491,378
No of ind.	53,248	53,248

Notes. All regressions are run with ln(exp+1) as the dependent variable and using the 95-trimmed sample of 53,248 migrants over years 2007-2016. In Model 1 the regression is run without independent variables of individual characteristics. Independent variables in Model 2 include indicators for age-gender group, individual income, marital status and number of children in the household.

5.3 Heterogeneous effects at the introduction of out-of-pocket prices

In Paper III, results show that at the introduction of out-of-pocket prices at age 20 in Region Västra Götaland, young adults decrease the number of visits to primary care physician by 7.1% on average, statistically significant at 1% level (Figure 5). As a comparison, the estimated effect for visits to specialist was a decrease of 2.6%, however significant only at 10% level (results not shown here). In subgroup analyses based on income quartiles, the results show that low income groups are more price sensitive compared with the subsequent quartiles (Figure 6). The number of visits to primary care physician decrease by 11.4% in the first (lowest) income quartile, 7.9% in the second, 5.5% in the third and 1.9% in the fourth. Results are statistically significant at 1% level for the first and the second quartile only. The results also show notable discrepancies based on sex (Figure 7). At the out-of-pocket price introduction, women reduce the number of visits by 9.2% (significant at 1% level) and men by 3.5% (statistically insignificant).

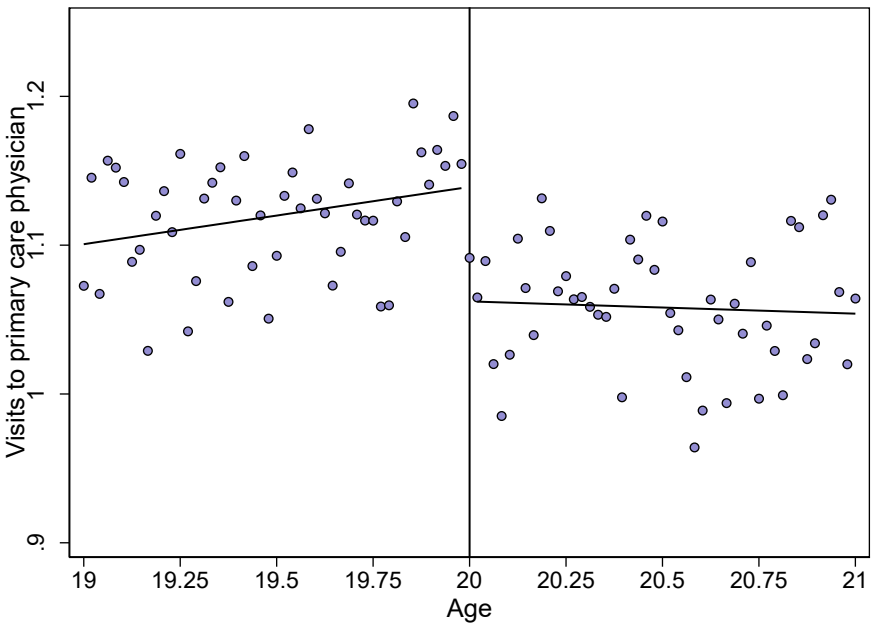

Figure 5. Main results of the introduction of out-of-pocket prices at age 20
Notes. The outcome unit is the number of visits to physician in primary care, per capita per year, in Västra Götaland. The dots show the average number of visits by age, the black line is the fitted line from equation (5), specified as linear splines. Adapted from Paper III (Johansson et al. 2019).

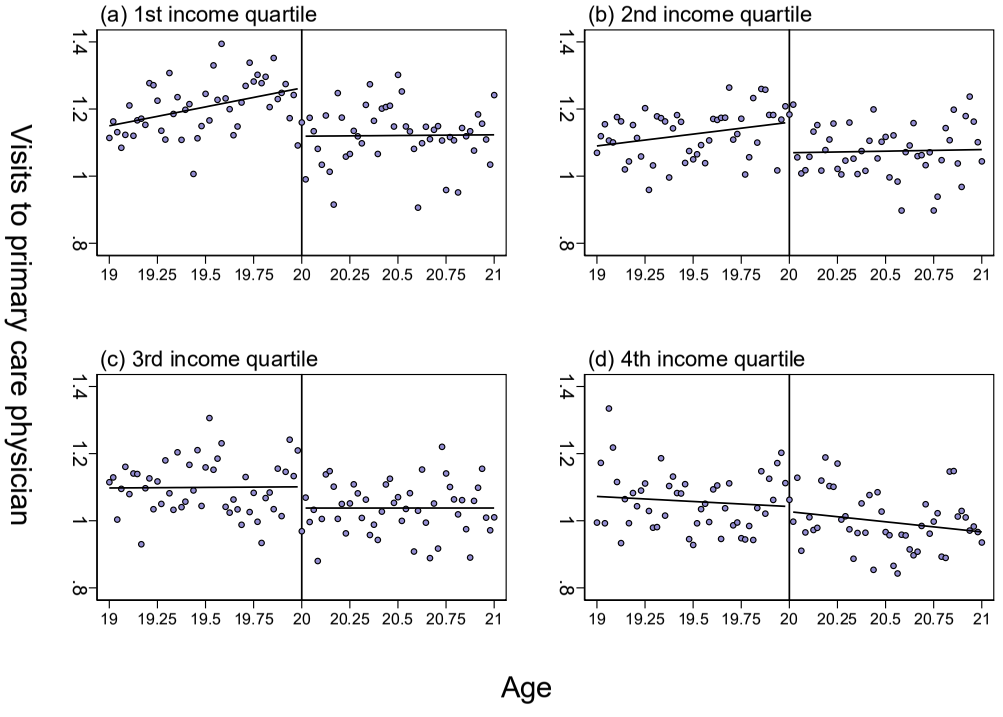

Figure 6. Discrepancies between income groups at the introduction of out-of-pocket price

Notes. Income is measured as the equivalized household income. The 1st income quartile refers to the group with lowest incomes. The outcome unit is the number of visits to physician in primary care, per capita per year, in Västra Götaland. The dots show the average number of visits by age, the black line is the fitted line from equation. Adapted from Paper III (Johansson et al. 2019).

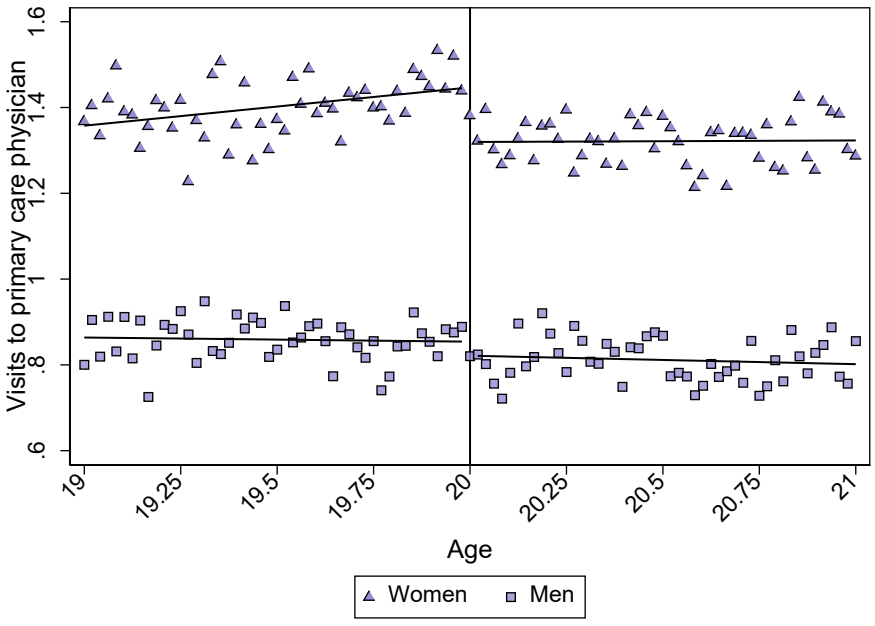

Figure 7. Discrepancies in by sex at the introduction of out-of-pocket prices

Notes. The outcome unit is the number of visits to physician in primary care, per capita per year, in Västra Götaland. The dots show the average number of visits by age, the black line is the fitted line from equation (5). Adapted from Paper III (Johansson et al. 2019).

5.4 Forward-looking behavior in the elimination of out-of-pocket prices

The results in Paper IV show that patients approaching the threshold of out-of-pocket price elimination at age 85, starts to delay primary care visits in four months (Region Stockholm) and two months (Region Västra Götaland) before the threshold, estimated by a statistically significant kink in the pre-policy underlying trend (Figure 8). The findings are evidence of forward-looking behavior and a delay effect among older adults. When facing a more generous cost sharing policy in the near future, older adults are willing to delay primary care visits. The delay effect corresponds to a 2.2% reduction of primary care visits in Stockholm and a 3.4% reduction of visits in Västra Götaland. In Stockholm, the delayed visits are shifted until after the policy threshold, while such an effect is not found in Västra Götaland.

Subcategorizing primary care visits by type of health care professional, the results show that the delay in visits in driven by a reduction in visits to non-physicians, true for both Stockholm and Västra Götaland (Figure 9). The immediate increase in visits after the threshold found in Stockholm, is driven by an increase in physician visits. The results show no evidence of a pure, persistent price effect of increased visits after the elimination of out-of-pocket prices.

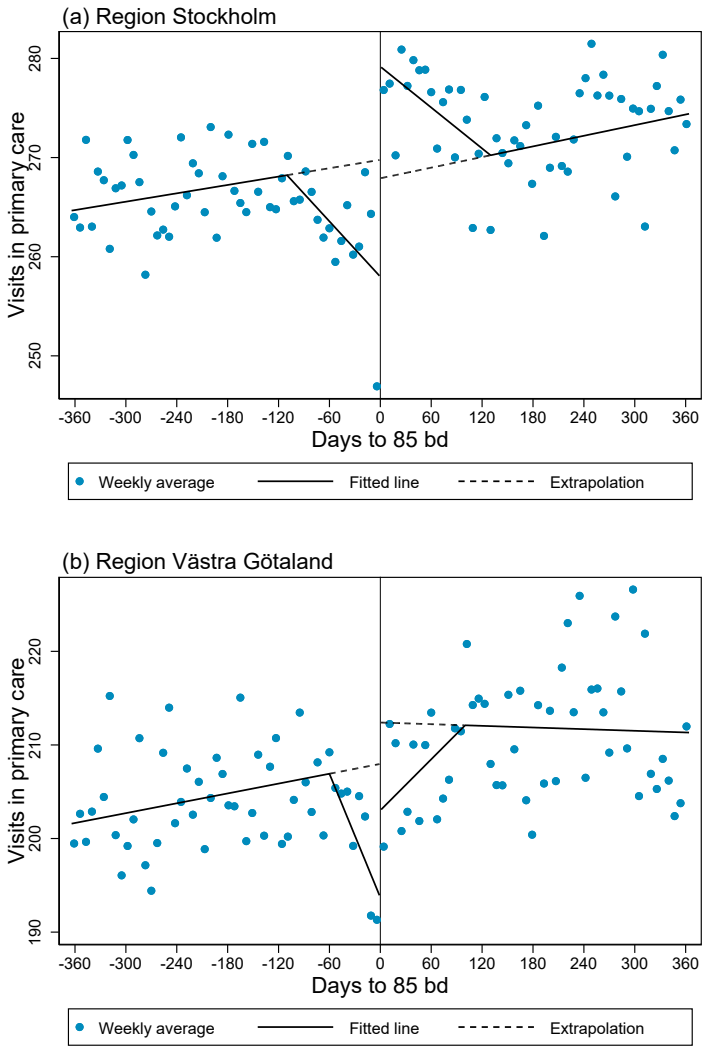

Figure 8. Main results of the out-of-pocket price elimination at age 85

Notes. Regression results estimated using equation (6). The outcome unit is primary care visits per day per 10,000 people. The blue dots are the weekly average of number of visits per day adjusted for a linear time trend, and the black line shows the fitted regression line. Figure adapted from Paper IV.

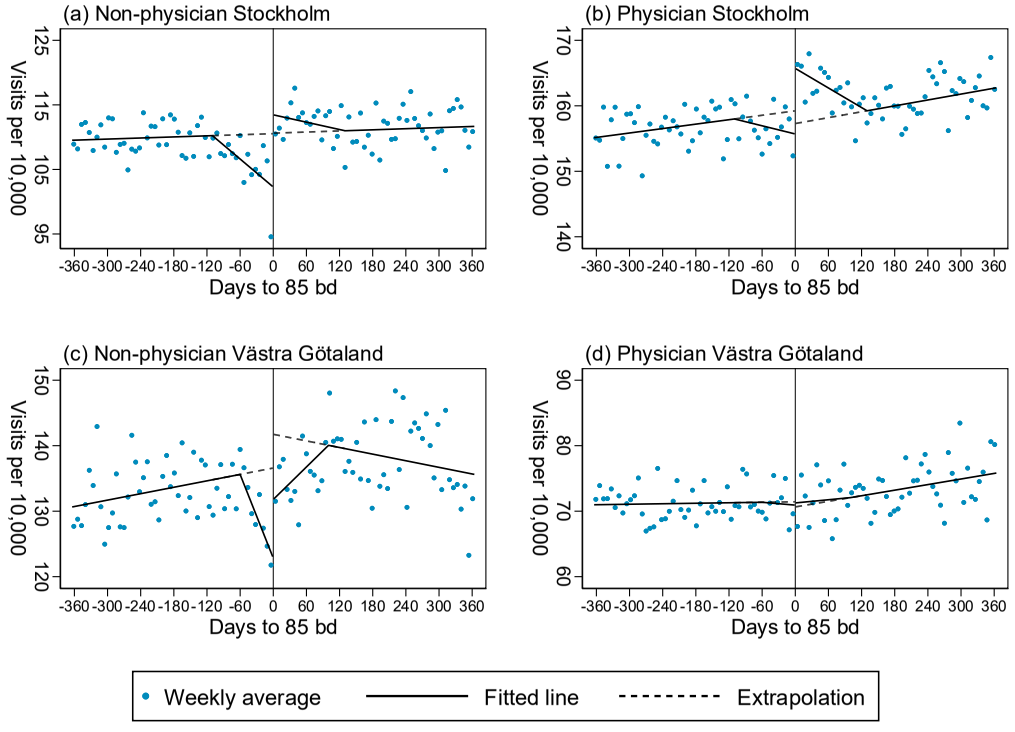

Figure 9. Subcategorized primary care visits at the elimination of out-of-pocket prices

Notes. Regression results estimated using equation (6). The outcome unit is visits to non-physicians in primary care for (a) & (c), and visits to physicians in primary care for (b) & (d). Please note that the y-axis differ in each graph. The blue dots are the weekly average of number of visits per day adjusted for a linear time trend, and the black line shows the fitted regression line. Figure adapted from Paper IV.

6 DISCUSSION

The central theme for this thesis is determinants of health care utilization with the purpose to contribute with new evidence on the two topics regional variation and price sensitivity in health care. Both topics have attracted attention in the scientific literature and this chapter begins with a thorough discussion of how the findings from the thesis can be interpreted and understood in relation to the previous literature. Thereafter, some methodological issues, ethical considerations and policy implications are discussed.

6.1 Determinants of regional variation

6.1.1 The size of regional variation differs by setting and by type of care

Given the current evidence, there is likely not one answer to the question of the determinants of regional variation but rather different answers depending on health care setting, type of health care and the level of geographical units. One of the first things to draw attention to is that the size of regional variation in health care varies in different countries and contexts. Descriptive statistics provided in this thesis show that regional variation in total health care expenditures in Sweden is relatively small, as the top-spending region has per capita expenditures about 18% higher than the lowest spending region (max/min ratio 1.18 in Table 1). The comparable figure in the Netherlands was 24%, in Germany 45%, in British Columbia, Canada 50% and in US Medicare about 200% (Fisher et al. 2009, Göpffarth et al. 2016, Lavergne et al. 2016, Moura et al. 2019).

The descriptive statistics in this thesis reveal also that the size of regional variation differ depending on type of health care. The number of physician visits was 58% higher in the top-utilizing region compared with the lowest utilization region, and for pharmaceutical expenditures the difference was 19% (Table 1). Measured by the coefficient of variation the size of regional variation was 0.12 for physician visits and 0.04 for pharmaceutical expenditures. Similar statistics have been shown in German data, where variation was larger for visits to psychotherapists compared with visits to primary care physicians and specialists, and in US Medicare, where regional variation for non-drug medical expenditures was higher compared with pharmaceutical expenditures (Kopetsch and Schmitz 2014, Zhang et al. 2010a). Descriptive statistics collected in a multinational report provided evidence of how the size of regional variation differ across a number of countries

and by various surgical procedures and hospital admissions (OECD 2014). Taken together, the descriptive evidence suggests that the institutional settings of the health care system, or even the specific policy setting by type of care within the same health care system, are key in understanding regional variation.

6.1.2 Two different perspectives on determinants of regional variation

The two studies on regional variation included in this thesis, set out in two different directions from the overarching question on the determinants of regional variation. In Paper I, we study the association between the average number of physician visits and regional level mortality, demographic factors and socioeconomic factors. This approach poses the question of whether regional variation can be explained by differences in medical need and preferences, which ties to the theoretical discussion of whether regional variation in health care can be justified for certain reasons, or if all regional variation is unwarranted (Skinner 2011). In the empirical literature differences in medical need or in population health have been seen as a justified reason for regional variation (see e.g. Göpffarth et al. 2016).

In Paper II, we use regional migration to separate the relative effect of individuals and of place (region) on regional variation. This approach takes more of a policy perspective on regional variation, not asking if it is justified, but rather how to deal with and reduce regional variation. If regional variation is mainly driven by differences in individuals' characteristics and preferences, policies with aim to change the institutional setting or other place-specific characteristics would have little impact on regional variation, or even be counterproductive (Finkelstein et al. 2016). These two approaches complement each other, the former estimating the degree of explanation of specific variables on regional variation and the latter separating the drivers of regional variation into an individual "demand-side" effect and a place-specific "supply-side" effect.

6.1.3 Individual effect and place effect

The main finding from Paper II is that about 9% of regional variation in pharmaceutical expenditures in Sweden is driven by a region effect and the rest about 91% by an individual effect. The results vary slightly depending on model specification and what groups of regions are being compared, but all show an overwhelming overweight to the individual effect and suggests that there is little scope for place-specific factors having caused regional variation in pharmaceutical expenditures in Sweden. Previous findings using the same empirical method have found a 50–60% place effect for health care utilization in US Medicare, a 50% place effect for hospital expenditures in Norway, a 30% place effect in total health

care expenditures in the Netherlands, and a 10% place effect for outpatient services in Germany (Finkelstein et al. 2016, Godøy and Huitfeldt 2020, Moura et al. 2019, Salm and Wübker 2020). Comparing these results to each other, it is important to note that besides comparing different countries, each study has used a different measure of outcome and type of care. Thus, one compares not only on a general level the institutional settings within each health care system, but rather the specific policy setting of for example outpatient services in Germany or prescribed pharmaceuticals in Sweden.

Salm and Wübker (2020) argued that high restrictions on physicians' practices and few constraints for patients may explain the limited place effect in regional variation in outpatient services in Germany. Considering the policy setting for prescribed pharmaceuticals in Sweden, our findings are plausible. The government authority TLV regulates what pharmaceuticals will be subsidized nationally, pharmacies are obliged to sell the cheapest alternative when generics are available and there are limited financial incentives for physicians to prescribe more or more expensive drugs. These kind of regulations are likely to limit the scope for a region effect. A potential mechanism of a place effect would be if high competition in primary care in some regions made physicians prescribe more pharmaceuticals to keep patients happy, however our results suggest that might not be the case here. The region-specific characteristics captured by the place effect are not necessarily only institutional settings, but other factors may as well be captured in the place effect. Godøy and Huitfeldt (2020) argued that part of the relatively large place effect in variation in Norwegian hospital expenditures may be related to the specific geographic and environmental conditions in Norway that make for example travel times long. Geographical conditions such as mountains and fiords are difficult to change, and are perhaps also an example of justified regional variation that is unaffected by individuals' medical need.

The findings from Paper II together with previous evidence estimating the relative effect of individuals and of place show clearly that the individual effect drives 40% or more of regional variation in all studied health care settings, for various types of health care. This result stands in contrast to a large part of the American literature on regional variation concluding that the "supply-side" is the main driver of regional variation in health care (Chandra et al. 2011, Fisher et al. 2009, Skinner 2011). It is of course plausible that the regional variation in US Medicare is to a large extent driven by supply-side factors such as physicians' preferences and financial incentives, but we need to be careful not to generalize such findings to completely diverse health care settings.

6.1.4 Determinants differ by type of care

One of the main findings in Paper I is that the factors that explain regional variation in physician visits differ by type of physician. Mortality and demography explain about half of regional variation in specialist visits, but has little association with primary care physician visits. Socioeconomic variables and variables of health care resources does not add to the degree of explanation in specialist care, but explains about one third of regional variation in primary care. These results show that the determinants of regional variation differ by type of care, within the same health care system. In the Swedish context, this may be related to differential incentives and regulations for primary and specialized care. A previous German study with similar empirical approach found that health and demography explained 45% of variation in primary care physician visits and 52% of variation in specialist visits (Kopetsch and Schmitz 2014). Diverse results by type of care has been found also in studies estimating the relative effect of individuals and of place. Moura et al. (2019) found a relatively smaller place effect for primary care expenditures and higher for pharmaceutical expenditures in the Netherlands, and Salm and Wübker (2020) showed that the place effect was relatively lower for primary care compared with specialist care in Germany. Despite this, differences in regional variation by type of care has been given little attention in the literature and the use of various outcome measures seems to be more of a coincidence than a deliberate choice, or simply based on data availability. It would be interesting to study further how the specific policy settings for various types of care within a health care system impact regional variation.

6.1.5 Large proportion of variation unexplained

The other main finding from Paper I is that 50–67% of regional variation in physician visits remain unexplained by included covariates. These results raise the question of what other factors may be explaining the rest of regional variation. From a statistical point of view, it is in practice impossible to explain 100% of variation. With that said, included variables explain but one third of variation in primary care visits and it is unlikely that the rest two thirds are just random variation. A previous German study with comparable empirical approach found that health, measured by life expectancy and a morbidity index, and demography explained about half of state-level variation in primary care physician visits and in specialist visits, and up to 70% when including structural, socioeconomic variables (Kopetsch and Schmitz 2014). Part of the explanation for the different findings may be that we do not fully control for health or medical need, while the German morbidity index does better in that aspect. Some US studies found that a minor share of regional variation in health care expenditures was explained by various measures of patients' characteristics and preferences, and having adjusted for the demand-side concluded that the supply-side is the main driver of regional

variation (Anthony et al. 2009, Baker et al. 2014, Sutherland et al. 2009). In Paper I, we cannot claim that we have completely adjusted for the demand-side, but rather likely suffer from omitted variable bias.

To be able to fully understand the specific factors that explain regional variation in physician visits a more precise measure of health or medical need would be preferable. However, even with a better measure of medical need, it is problematic to use as an independent variable, because of health's endogenous relation to health care utilization (expenditures). As mentioned previously, regional variation in health care may be considered justified if caused by differences in medical need – but in the long run one would expect higher levels of health care utilization to lead to improved health. How can we estimate the effect of today's health as a determinant for today's regional variation in health care utilization, if today's health is in itself an effect of yesterday's health care utilization? This is a complex methodological issue.

There may be other place-specific characteristics and institutional settings that were omitted from the analysis in Paper I that play an important part in understanding regional variation in physician visits. In the decentralized Swedish health care system, each region set policy regulations and incentives for physicians and health care providers that likely play a part in the level of health care utilization. The location of universities and university hospitals is another factor that may affect regional variation, both from the demand-side and the supply-side. In regions with university hospitals where medical professionals are educated, the number of physicians and the degree of medical specialization may affect the use of both primary and specialized services. The location of universities is correlated also with the average educational level and the age structure of the region, which in turn may impact health care need as higher education and young age usually is associated with better health.

6.2 The effects of out-of-pocket prices on primary care use

6.2.1 Are young more price sensitive than older adults?

With respect to price sensitivity, one of the main findings from Paper III was that young adults reduced the number of physician visits in primary care at the introduction of patient out-of-pocket prices at age 20. These results are in line with a large part of the previous empirical literature documenting that price of health care causally impact the level of health care use (e.g. Aron-Dine et al. 2013, Einav and Finkelstein 2018). Our results specifically relate to and are consistent

with findings from similar studies showing how children, adolescents and young adults respond to an introduction of out-of-pocket prices (Han et al. 2019, Nilsson and Paul 2018, Olsen and Melberg 2018, Remmerswaal et al. 2019a).

In contrast, in Paper IV we did not find evidence of a permanent increase in the number of primary care visits following the elimination of patient out-of-pocket prices at age 85. These results diverge from previous literature with focus on older adults in the US and in Japan, where insurance eligibility or reduced out-of-pocket prices for older adults led to increased use of health services (Card et al. 2008, Fukushima et al. 2016, Shigeoka 2014). The different results may possibly be explained by a lower price sensitivity among older adults (discussed below) and diverse health care settings where the Swedish system have relatively low out-of-pocket prices. In Paper IV, we assess a reduction in out-of-pocket prices from a low level (small relative to incomes) to zero (free-of-charge). Card et al. (2008) studied insurance eligibility in the US, with relatively high out-of-pocket prices and in some cases a large difference in out-of-pocket price before and after eligibility. Shigeoka (2014) noted that out-of-pocket payments in Japan can create considerable financial burden for older adults, as out-of-pocket costs for inpatient care can account for up to 27% of the average monthly income.

It is plausible that individuals are less price sensitive with respect to relatively lower out-of-pocket prices, since they are less likely to be affected by liquidity constraints. This interpretation is supported by evidence from previous studies of the general population in Sweden and in Germany that did not find evidence of a response in the number of physician visits in the general population following a change in out-of-pocket prices around €10–20 (Jakobsson and Svensson 2016b, Kunz and Winkelmann 2017, Schreyögg and Grabka 2010).

The findings from this thesis suggest a heterogeneity in price sensitivity with regards to age – that young adults seem to be more price sensitive in health care compared with older adults. Previous evidence is limited, but in studies where direct comparison across age groups have been made, Farbmacher and Winter (2013) and Simonsen et al. (2016) came to the same conclusion with respect to physician visits in Germany and pharmaceutical expenditures in Denmark. Hayen et al. (2018), on the other hand, did not find evidence of heterogeneity in price sensitivity across age groups in total health care expenditures in the Netherlands. The differences between mentioned studies is likely related to the different types of health care assessed, what age groups are compared, or potentially cultural differences across the European countries.

A plausible reason for older adults to be less responsive to patient out-of-pocket prices compared with young adults is differences in health status. There is

(obviously) a strong correlation between age and health, especially comparing 20-year-olds to 85-year-olds as in this thesis, and with age comes an increased risk for chronic diseases and high severity diseases, and thus an increased need for health care. It is reasonable to assume that an individual in their decision to seek or not to seek health care, is less concerned about the out-of-pocket price for a health issue of higher severity. Some previous studies have shown that individuals with chronic illness are less price sensitive compared with healthier individuals (Chandra et al. 2014, Fukushima et al. 2016), but the evidence is limited. With current evidence, it is difficult to say whether the diverse findings of the different age groups is driven by differences in health status, if there are other correlated factors or if age itself has a causal effect on price sensitivity.

6.2.2 Heterogeneity based on type of care, income, and sex

In Paper III and IV, we find that patients respond less to the out-of-pocket price for specialist visits compared with visits to primary care physician, and less to physician visits compared with nurse visits in primary care. These results are in line with the above proposed hypothesis that the severity of illness may affect individuals' responsiveness to price. In the previous literature, the results are mixed but some studies have shown that patients were less price sensitive to inpatient care compared with outpatient care and that patient were less price sensitive to specialized care compared with primary care (Chandra et al. 2014, Finkelstein et al. 2012, Han et al. 2019, Nilsson and Paul 2018, Shigeoka 2014). The price responsiveness by type of health care likely differ not only depending on severity of illness but also the amount of out-of-pocket price and institutional settings such as need for referrals and waiting times.

One of the contributions of Paper III, is that we can credibly show that low-income groups are more price sensitive compared with high-income groups. Theoretically, it is reasonable to assume that low-income groups are more sensitive to prices (Baicker and Goldman 2011). Our results are consistent with findings from the only previous study, to my knowledge, that have made direct comparisons by income level with income data on individual level (Nilsson and Paul 2018). Similar to Nilsson and Paul (2018), who used data from the Swedish Region Skåne, the data in Paper III cover the entire population of the age group studied thus including the full spectra of the income distribution. Additionally, the Swedish universal health care system implies that very few individuals use health care services outside of the public system. These features enable an (unusual) direct comparison of price sensitivity in health care across income groups by limiting potential selection effects.

The results in Paper III show that women are more price sensitive compared with men, and the evidence from the previous literature is mixed (Farbmacher and Winter 2013, Hayen et al. 2018, Olsen and Melberg 2018). It is difficult to intuitively explain why either men or women would be more price sensitive, but related factors could be income, and social norms in health care seeking behavior. It has been shown that women in Sweden use more health care services compared with men, especially primary care (Osika Friberg et al. 2016).

6.2.3 Forward-looking moral hazard

Despite the lack of evidence of a permanent increase in primary care visits after the elimination of out-of-pocket prices at the 85th birthday, the findings from Paper IV show that older adults are price sensitive in a forward-looking manner. In the months before the 85th birthday, individuals delayed primary care visits and in Region Stockholm the visits were shifted until after the birthday. The findings correspond to previous research documenting forward-looking behavior with respect to out-of-pocket prices (Aron-Dine et al. 2015, Klein et al. 2020). A few previous studies have shown delay of pharmaceutical consumption awaiting a more generous policy or an insurance contract's annual reset (Alpert 2016, Einav et al. 2015). But our findings are the first, to my knowledge, to provide evidence of a delay effect in primary care visits in response to an age threshold in the out-of-pocket price scheme. Additionally, the findings suggest that patients delay non-physician visits rather than physician visits, perhaps because those visits and health issues are considered less urgent or less severe. Similarly, Alpert (2016) showed that older adults delayed drugs for chronic diseases but not drugs for acute events.

The conflicting results in Paper IV between Region Stockholm and Region Västra Götaland may be explained by differential amounts of out-of-pocket prices and differential financial incentives for the health care providers in the two regions. Out-of-pocket prices are higher in Stockholm compared with Västra Götaland, which implies that the benefit of delaying a health care visit is higher in Stockholm and a longer period of delay is plausible. The provider reimbursement in Stockholm is in part based on fee-for-service, while in Västra Götaland the provider reimbursement is only based on risk-adjusted capitation. Thus, health care providers in Stockholm has a financial incentive to replace a delayed visit, and may encourage price sensitive patients to return after the 85th birthday instead.

In Paper III, there is limited evidence of forward-looking behavior among young adults with respect to the upcoming introduction of out-of-pocket prices at the 20th birthday. Likewise, previous studies assessing price sensitivity among children, adolescents and young adults tested but did not find evidence of

forward-looking behavior (Han et al. 2019, Nilsson and Paul 2018, Remmerswaal et al. 2019a). One possible explanation for these results is that young are less forward-looking with respect to patient out-of-pocket prices than older adults. On the other hand, considering the introduction of out-of-pocket prices at a certain age threshold, a forward-looking individual would need to bring forward health care visits as opposed to delay visits, which may be more difficult. The current evidence does not provide the answer to these somewhat conflicting results, and the issue of forward-looking moral hazard will need further research.

6.3 Methodological issues

6.3.1 Describing regional variation

Studying regional variation, there are a number of methodological choices to be made, for example regarding the measure of outcome, how to describe the size of variation, and geographical units; and the previous literature contains various examples. First, one may ask whether health care expenditures or health care utilization should be used as the measure of outcome. While it in practice may be a question of data availability, there may in theory be advantages and disadvantages with both measures. From an equity perspective, and assuming that regional variation is by some means justified, utilization may be preferred as the measure of outcome because it eliminates differences in expenditures that might occur for other acceptable reasons. For example, geographical or environmental conditions, possibly related to population density, may lead to higher expenditures that are unrelated to medical need but may still be considered justified. As mentioned above, Godøy and Huitfeldt (2020) discuss in their article that long travel times due to geographical conditions in Norway may be an explanation for a strong region effect in variation in hospital expenditures. Similarly, Lavergne et al. (2016) consider geographical conditions in Canada as an acceptable reason for regional variation in health care expenditures.

On the other hand, the descriptive statistics provided in this thesis (Table 1) show that regional variation is larger in utilization rates (physician visits) compared with rates of expenditures (total health care). That raises the question of efficiency – are some regions more efficient in their use of resources and manage higher rates of visits to the same level of expenditures? To assess if regional variation is a sign of inefficiency, expenditures may be preferable as the measure of outcome. Another advantage of using expenditures is that you can get one summarized measure of all types of health care; visits, admissions, pharmaceuticals; and with the option to subcategorize.

Second, to describe regional variation in a certain country, a map and other graphical tools are of course recommended, but a summarizing measure is needed to compare across countries. The ratio between the top and bottom region (essentially the percentage difference) is a straightforward and intuitive measure of differences between regions. However, one may ask which is a more representative measure: the ratio of maximum to minimum, the 90th to the 10th percentile, or the top and bottom quartile? The coefficient of variation on the other hand, defined as the ratio of the standard deviation to the mean, propose a measure more representative of the overall regional variation, but is less intuitive. There is need for caution when comparing regional variation across countries, to make sure that one does not compare apples with oranges.

Third, the level of geographical units used to assess regional variation will certainly affect the analysis, but the issue has received limited attention in the literature. Previous evidence provides examples of regional variation across states, provinces, counties, municipalities, hospital referral regions, hospital service areas, and post-codes areas. The variation is (by definition) higher when comparing a larger number of smaller units (OECD 2014). Zhang et al. (2012) compared regional variation in hospital referral regions to variation in local hospital services areas in US Medicare and concluded that hospital referral regions may be too crude as there may be relevant differences within the regions. Göpffarth et al. (2016) showed that the level of explanation of included variables differed depending on whether variation was assessed on state-level or county-level. In the studies included in this thesis, variation is assessed on regional level, because the regions are the organizational units in the Swedish health care system and because of data availability. It would be interesting to study regional variation in Sweden on for example municipality level or Statistics Sweden's demographic statistical areas (DeSO), to assess the variation within regions and what may explain regional variation on a lower level of aggregation.

6.3.2 Comparison of the event study and the decomposition analysis

In Paper II, estimating the relative effect of region and of individuals on regional variation in pharmaceutical expenditures, the event study and the decomposition analysis provide slightly differing results. The empirical approach, based on regional migration and two different types of fixed effects models, is methodologically complex and the comparison between the models is not completely straightforward. Both models' identification strategy builds on patient migration, as moving from one region to another creates a quasi-experimental setting where individuals who move find themselves, from one year to another, in a different region with (potentially) diverse characteristics that may (or may not)

affect their level of health care expenditures (utilization). Using migration to assess how place-specific characteristics impact individuals' development is not new (Song et al. 2010).

The event study with a two-way fixed effects model is quite intuitive. The main independent variable is constructed by the average difference between region of origin and region of destination and estimates how the regional differences (δ_i) in average expenditures (utilization) affect individual expenditures (utilization) of a regional migrant. The estimated region effect may be interpreted causally if the independent variables, including the variable for the average difference between regions (δ_i), are unrelated to unobserved, time-varying individual level characteristics captured in the error term (Salm and Wübker 2020). Potential violations to the described exogeneity assumption would be if regions with higher health expenditures is more attractive to unhealthy people, if the region effect varies over time and if there is heterogeneity in the region effect depending on the type of region (Salm and Wübker 2020).

Salm and Wübker (2020) proposed testing for these violations, and a limitation in Paper II is that we have not performed such tests, which would improve the causal interpretation of our results. Another limitation to raise regarding the event study is how representative the regional migrants are to the whole population, and whether we can assume that those who stay are affected the same way by regional characteristics. The regional migrants are younger, higher educated and fewer are married than the stayers (Table 2 in Paper II). It is likely that the region effect of stayers' health expenditures may be larger or smaller than the region effect of the migrants. A proposed way to deal with these differences is to use propensity score matching or weighting to find a sample of migrants that are more like the general population and test if their estimated region effect deviates from the pure sample of migrants.

In the decomposition analysis with a three-way fixed effects model, the estimated coefficients of the regression are used in a decomposition exercise, as opposed to evaluating the coefficients directly. An advantage of the decomposition analysis is that the full sample of migrants and stayers is used, and the individuals who stay in their home region function as the counterfactual of how health care expenditures may have altered over time even without moving. In the decomposition, we calculate average expenditures in a group of regions (e.g. the top half of regions) to compare with another group of regions (the bottom half), which gives the average difference in health expenditures between the two groups. With the estimated region fixed effects coefficients, we calculate the average region fixed effect in each group and the average difference in region fixed effect between the two groups.

In a sense, the event study regression takes into account all possible combinations of region pairs (of origin and destination region), and regions with more frequent migrations will have a higher weight to the estimated region effect. A (potential) disadvantage of the decomposition exercise is that it depends crucially on what groups of regions are compared, and included regions are weighted the same irrespective of size, population or number of migrants. It would be interesting to take such measures into account, for example by weighing regions by size, to assess how the results of the decomposition may be affected. Moura et al. (2019) and Salm and Wübker (2020) showed that the region effect estimated in the decomposition analysis varied considerably when comparing regions based on for example high and low proportion of older adults, as opposed to compare high and low spending (utilizing) regions. In summary, perhaps one should not be too specific with the exact number of the estimated region and individual effects when comparing the results of the event study and of the decomposition analysis. The question that we may answer in these analyses is rather whether the determinants of regional variation primarily stems from place-specific or individual level characteristics. The findings from Paper II of both event study and decomposition clearly shows that the individual effect is outweighing the place effect in regional variation in pharmaceutical expenditures.

6.3.3 Causal interpretation of the RD design

Regression discontinuity (RD) design, the base for the analyses in Paper III and IV, is proposed as way to estimate the causal effect of a policy treatment, under certain assumptions. The RD design is built on a policy setting where an assignment variable, in this case age, strictly determines the treatment, in this case the patient out-of-pocket price (Angrist and Pischke 2014, Lee and Lemieux 2010). Theoretically, the policy rule regulates the threshold value (at what age out-of-pocket prices are introduced or eliminated) and we assume that individuals above the threshold are treated, but there may in practice be situations that deviates from the policy rule. Such a situation would arise if health care professionals, for example an altruistic reception nurse, let the patient skip the out-of-pocket payment or make sure to schedule a visits before (after) the threshold to the benefit of the patient.

In the basic design of the RD model we want to get as close as possible to the threshold, to credibly assume that individuals of age 19 years and 360 days are essentially no different from individuals who are 20 years and 7 days, almost as if randomly assigned by their birthday to either side of the threshold (Lee and Lemieux 2010). A potential violation to this assumption is if we for some reason believe that individuals are not "randomly distributed" over the threshold. In the two studies included in this thesis, we follow individuals over time, so essentially

we observe their health care visits during one year before and one year after the threshold. In Paper III however, the data set is limited to two years (2014–2015), which implies that each individual cannot be followed from their 19th birthday to their 21st birthday. This creates a "rolling sample" over the age period studied, which may violate the causal interpretation, but in the weeks right before and right after the 20th birthday, there is likely a very limited effect of the rolling sample.

In Paper IV, the data set covers a longer time period (2014–2018) so most individuals may be followed from their 84th birthday to their 86th birthday. Instead in this sample, mortality poses a threat since almost 30% of the sample passes away during the time period. Thus, there is need to consider how to handle the people who pass away. One way would be to exclude them from the analysis, but we have chosen to include them and censor them for the time after their death. It may be that individuals in the last months before death use more outpatient services, or less outpatient services because they are admitted to hospital. Anyhow, because the individuals do not pass away discontinuously at their 85th birthday, mortality does not formally pose a problem for the causal interpretation of the estimated effect at the threshold.

The causal interpretation of the estimated effect at the threshold rests on the assumption that no other factors that may affect the outcome of interest change discretely at the threshold (Angrist and Pischke 2014, Lee and Lemieux 2010). If there are such factors, the estimated effect should be interpreted as the total effect of all factors that change discretely at the threshold. To my knowledge, the 85th birthday does not serve as a threshold for other policy rules or health related matters such as health check-ups or need to prove medical ability for driving license.

Around the age of 20, on the other hand, there are many life developing changes such as leaving home and staring a new job or education that may also impact health and health care utilization, but these are continuous changes over time, not specifically at everyone's 20th birthday. But there is in the Swedish context another relevant public health policy rule: the legal age for drinking alcohol. Previous research has shown that the legal drinking age in the US lead to a discrete increase in mortality (Carpenter and Dobkin 2009). It is plausible that the policy age threshold for alcohol in Sweden may lead to worse health and a discontinuous increase in the use of health care services, and in that case, the estimated reduction in visits at the introduction of copayments at age 20 would be an underestimate of the true price effect. However, the dependent variable in Paper III is visits to primary care physician, while sudden alcohol-related health problems such as vehicle accidents or violence likely has more effect on emergency visits. Besides, the Swedish policy for the legal drinking age contain two thresholds: permission

to drink alcohol in restaurants and bars at age 18 and permission to buy alcohol in liquor stores at age 20. A policy which may smooth out some of the negative effects of the legal drinking age.

The causal interpretation of the RD design also rests on the assumption of no manipulation around the threshold (Angrist and Pischke 2014, Lee and Lemieux 2010). In general, this may relate to tendencies of individuals choosing a value just below or just above the threshold. A related example is the distribution of pharmaceutical expenditures among individuals covered by US Medicare, where a disproportionately large number of individuals end up with annual expenditures just at the threshold for increased out-of-pocket prices (Einav et al. 2015). In the case of age and out-of-pocket prices, it seems unlikely that individuals would manipulate their age to avoid an (inevitable) threshold at age 20 or age 85. Rather, there may be manipulation in the timing of health care visits if individuals are forward-looking and anticipate the upcoming policy change.

We deal with this kind of manipulation in Paper III by using a donut regression to test whether there is any evidence of forward-looking behavior, and in Paper IV by developing a donut RD with kink regression to specifically estimate the scope of forward-looking behavior. Finding evidence of forward-looking behavior and manipulation in the timing of primary care visits in Paper IV, is in a sense a violation of the causal interpretation of the RD model. Additionally, the idea behind the donut regression is to step away from the threshold (exclude observations close to the threshold) and that too raises the question of whether the causal interpretation of the estimated effect still is valid. Close to the threshold we may assume, as mentioned above, that individuals just days before the 20th birthday are no different from those in days just after the 20th birthday. But extending the gap to for example 84 years and 9 months compared with 85 years and 3 months, there may be systematic differences in health and health care need between those six months, especially considering the older adults (probably less so with the 20-year-olds). In the donut RD kink model in Paper IV, we estimate kink points two to four months away from the threshold, and interpret these kink points as an effect of the copayment elimination. Although we should be careful to interpret the kink points as causal, it is difficult to think of other factors than forward-looking behavior that would make 84-years-and-9-months-old individuals reduce the number of primary care visits in the months before the copayment elimination.

6.3.4 The out-of-pocket limit

One of the limitations in Paper III and IV on the impact of patient out-of-pocket prices is that we cannot account for the annual out-of-pocket limit in outpatient

care. In Paper III studying young adults around the age of 20, the out-of-pocket limit is unlikely to affect the results since the average number of visits in outpatient care is very low (about one primary care physician visit per year, Table 1 in Paper III) and few young adults reach the out-of-pocket limit. In Paper IV on the other hand, older adults have a much higher level of health care utilization with about 12 outpatient visits per year (including primary and specialist, physician and non-physician visits) so a large number of individuals are affected by the out-of-pocket limit (Table 3 in Paper IV). The out-of-pocket limit is potentially a reason why we find older adults are less price sensitive compared with young adults. It may also explain that our results differ from the American and Japanese studies of older adults' response to patient out-of-pocket prices; it is however unclear whether similar policies are present and/or accounted for in those studies (Card et al. 2008, Fukushima et al. 2016, Shigeoka 2014).

In Paper IV, 84-year-olds who had already reached the out-of-pocket limit in the months before their 85th birthday would in theory not be affected by the upcoming out-of-pocket price elimination because they already had access to free-of-charge health care. In that case, then the estimated delay effect that we find in the months before the 85th birthday is an underestimate of the true effect because part of the sample does not respond to the upcoming out-of-pocket price elimination. It is however difficult to assess this in practice. The design on the out-of-pocket limit as rolling over the year and some inconsistencies in the databases regarding registration of relevant variables make it difficult to determine when in time the individual actually have reached the out-of-pocket limit, and when her rolling year will reset. In robustness tests in Paper IV, we find that the results of the delay effect differ for paid visits and visits free-of-charge in the year before the 85th birthday. We find that in Västra Götaland paid visits, but not "free" visits, are delayed in the months before the out-of-pocket price elimination. In Stockholm however the results are inconclusive. In summary, it is likely that the out-of-pocket limit plays an important part in individuals' price sensitivity in health care and it would be very interesting to design a study to specifically study its effects on health care utilization.

6.4 Policy implications

Health policy targeting region-specific characteristics or institutional settings with aims to reduce regional variation in pharmaceutical expenditures in Sweden would have limited effect since, as the findings from this thesis shows, regional variation in pharmaceutical expenditures is primarily driven by differences in individual level characteristics. It is difficult to draw specific policy advice regarding regional variation in physician visits based on the findings in Paper I. Considering a

relatively larger regional variation in physician visits compared with pharmaceutical expenditures, it is important to establish a better understanding of the driving factors of regional variation in physician visits, both in primary care and in specialized care.

Policymakers should be aware that even relatively small out-of-pocket prices affect the level of primary care use, but with heterogeneous effects based on age, income and sex. Finding that some groups are more price sensitive is important with respect to equality in health and health care, and policymakers should consider what type of distributional effects are acceptable or unwarranted. In design of out-of-pocket price policies, policymakers need to consider forward-looking behavior with respect to patient out-of-pocket prices, and for example evaluate the risk of health care delays leading to worse health outcomes. Differential effects in different regions indicate that it is important to acknowledge the interaction between out-of-pocket price policies and other policy regulations and incentives.

The changes in out-of-pocket prices at age 20 and age 85 result in the following cost implications. The introduction of out-of-pocket prices for young adults is beneficial to the regional government as it results in a revenue from collected out-of-pocket payments and a reduction in health care spending due to decreased number of physician visits. The elimination of out-of-pocket prices for older adults reduces the financial burden for the patients, now carried by the national government, and without (direct) increased costs for the regional government (because we found no evidence of a permanent increase in health care visits). From a societal point of view, the collection and the exemption of out-of-pocket payments is not a cost but a transfer. The delay effect caused by the elimination of out-of-pocket prices for older adults reduce health care spending for the regional government, but in Stockholm the shift of the number of physician visits until after the 85th birthday implies an increased cost to the regional government.

6.5 Ethical considerations

The studies included in this thesis did not involve any physical procedures on human beings, and posed no risks to the health of the research subjects. As a researcher, one need to consider the ethical aspects of the risks that the research may have on research subjects' safety, integrity and human dignity, a risk which should be weighed against the academic and societal benefits of gaining new knowledge from research. The main risk associated with the studies included in this thesis is that individual level register data contain specific, sensitive and delicate information about a large number of individuals. For the purpose of this

thesis, data of health care utilization have been merged with demographic and socioeconomic information. This means that for example information about when and where health care visits have been made, information about diagnoses and prescribed pharmaceuticals, and information on individual income, birthdate and place of residence, are available in one large data set. The data has been anonymized and cleared of personal identification numbers, but it may, at least in theory, be possible to identify individuals in the data set. It becomes clear that such information in the wrong hands poses a risk to individuals' integrity and human dignity. Having said that, the risk that a third party (with vicious intentions) get access to the data is very small.

As a researcher, one of the first steps to protect individual integrity and dignity is to request only the data that is specifically needed for the proposed research project. Once data has been provided from the register holders, it is important to keep it stored safely with access only to the research group. For the research conducted in this thesis, the academic and societal benefits of new knowledge outweigh the potential risks to the research subjects' integrity and human dignity. As a last note, one may consider the use of register data from a different perspective: given the large amounts of detailed data collected and stored in Swedish register, it would be unethical and a waste not to use them for purposeful research with benefits to society.

7 CONCLUSION

The aim for this thesis was to contribute with new knowledge on the determinants of regional variation in health care and how individuals respond to patient out-of-pocket prices in Sweden. The findings from the thesis show that the factors explaining regional variation in outpatient physician visits differ in primary and specialized care, but that half or more variation in physician visits remain unexplained by included covariates. For regional variation in pharmaceutical expenditures, the results show that individual level characteristics outweigh, to about 90%, regional characteristics as the driving determinants of variation, which implies that health policy targeting place-specific characteristics would have limited impact on regional variation in pharmaceutical expenditures. The findings suggest that the specific institutional settings, which often differ by type of care within the same health care system, are key in understanding regional variation. Further studies are needed to fully understand both the specific factors and the relative effect of individual and of place in determining regional variation, for all types of health care.

Moreover, the findings from the thesis show that 20-year-olds reduce the number of visits to primary care physician by 7% when introduced to a patient out-of-pocket prices and that women and low-income groups are more price sensitive compared to their counterparts. For older adults approaching their 85th birthday and an elimination of out-of-pocket prices, the results show a delay of primary care visits by 2–3% in the months before the elimination, but without a persistent increase in the number of visits after the out-of-pocket price elimination. Despite relatively small effects, these results imply that patient out-of-pocket prices is an important determinant of health care utilization, even in a universal, highly subsidized health care system with relatively low out-of-pocket prices. Finding that the responsiveness to out-of-pocket prices varies in different Swedish regions suggests that there is in a sense a piece of regional variation in price sensitivity which may be explained by an interaction between out-of-pocket price policies and other health policy regulations. In conclusion, the thesis provides important evidence on forward-looking behavior and on heterogeneity in price sensitivity in health care with respect to age, income and sex, aspects that policymakers ought to take into account when developing health policy.

8 FUTURE PERSPECTIVES

Current evidence on regional variation in health care is inconclusive, which suggests that there is not one answer to the question of the determinants of regional variation, but that different factors may explain regional variation depending on the health care setting. The findings from this thesis indicate that even the specific institutional settings by type of care may be important. Future research should aim to assess regional variation by type of care within the same health care system, and to study in what way specific policy settings may enhance or discourage regional variation. This thesis provides evidence of how regional variation in pharmaceutical expenditures is primarily driven by individual level characteristics, but further research is needed to establish what specific individual level factors are the most important drivers, and which individual level characteristics may be seen as justified reasons for regional variation. A challenge that remains is how to fully adjust for differences in medical need, and to establish the causal relation between health and regional variation in health care utilization.

For future research on the determinants of regional variation, both the methodological approaches used in this thesis are relevant to establish the relative effect of individual and of place, and to study the impact of specific factors on regional variation. Variation has been documented across various geographical units such as regions, provinces and hospital referral regions, but there is need for more knowledge of what determines regional variation on different level of aggregation. In the Swedish setting, it would be interesting to compare determinants of regional variation for example across regions and across municipalities.

It is established in economic theory and in the empirical literature that higher out-of-pocket price for health care causally decrease the level of health care utilization. The findings from this thesis contribute with evidence that individuals are price sensitive with respect to health care also in a universal, highly subsidized health care system. However, the consequences of moral hazard and price sensitivity in health care are still uncertain, especially with respect to health outcomes in the long run. Future research should aim to determine what type of visits for what type of health issues are skipped (or delayed) due to patient out-of-pocket prices, and whether skipping health care lead to worse health outcomes in the long run and in turn increased costs of health care in the future.

The thesis provides evidence of heterogeneous effects of out-of-pocket prices based on income and sex, and indicate differential responses by age, and there is

need for more knowledge on the driving causes and mechanisms of heterogeneity in price sensitivity in health care, for example, whether differential responses between younger and older is an effect of age or of health. Additionally, the results show differential effects by region in the response to out-of-pocket prices, and it would be relevant to focus future research on how other policy regulations and incentives interact with price sensitivity in health care. The findings from the thesis contribute to the emerging literature on forward-looking moral hazard and further research is needed on the effects of various dynamic incentives such as by various types of out-of-pocket payments, by out-of-pocket limits and whether individuals respond differently depending on facing a reduction or an increase in out-of-pocket prices. It would also be interesting to study heterogeneity in forward-looking behavior in health care with respect to demographic and socioeconomic groups.

ACKNOWLEDGEMENT

I'd like to take the opportunity to express my gratitude to everyone who has supported me through this process.

My supervisor Mikael, thanks for giving me this opportunity, for your guidance, your feedback and your trust. You have never been further than an email or a video link away, even though we have often been in different cities, or continents. My co-supervisor Niklas, thanks for adding optimistic and constructive perspectives to research, and to life.

To my former office roomie Josefine, thanks for many long talks and for sharing your knowledge, your dedication is an inspiration to me. To my office roomie Frida, thanks for all good discussions high and low, and for motivating me to go orienteering again, I will miss our small talk at the desks. My colleagues at Health Metrics and AMM, and later on the School of Public Health and Community Medicine: Sara, Catrin, Anna, Adnan, Laith, Fredrik, and others, thanks for introducing me to academia and to Gothenburg.

A special thanks to my Australian coauthors of Paper IV: thanks Dennis, Sonja and Johannes for welcoming me to visit the Centre for Health Economics at Monash and for your much valuable thoughts, encouragement and feedback.

To all fantastic and inspiring PhD students I have connected with through these years, thank you! Many of whom I've met through the PhD Student Council at the Sahlgrenska Academy – Malin, Kajsa and Daniel in joint work with the PhD students' work environment survey; teaming up with Cristiana and Adnan in the Council presidium; and starting up the Doctoral Committee at the Institute of Medicine with Emma, Daniela, and Emanuele. My fellow health economists, Ellen and Sofie, it has been great to share the PhD journey with you. And to many other PhD students and researchers that I have met and got to know in Gothenburg, in Sweden and abroad. This has been one of the most important parts of my PhD experience, to create a network and a place to belong. Let's keep in touch!

To my family, thanks for believing in me. Hamed, thanks for pushing me, tu mera hero, main tera hero.

REFERENCES

Agha, L., B. Frandsen, and J.B. Rebitzer. 2019. Fragmented division of labor and healthcare costs: Evidence from moves across regions. *Journal of Public Economics*, 169: 144-159.

Ahammer, A., and T. Schober. 2020. Exploring variations in health-care expenditures—What is the role of practice styles? *Health Economics*, 29(6): 683-699.

Alpert, A. 2016. The anticipatory effects of Medicare Part D on drug utilization. *Journal of Health Economics*, 49: 28-45.

Anell, A. 2015. The public-private pendulum–patient choice and equity in Sweden. *New England Journal of Medicine*, 372(1): 1-4.

Anell, A., A.H. Glenngard, and S.M. Merkur. 2012. Sweden: Health system review. *Health systems in transition*, 14(5): 1-159.

Angrist, J.D., and J.-S. Pischke. 2014. *Mastering 'metrics: the path from cause to effect*. Princeton, New Jersey: Princeton University Press.

Anthony, D.L., M.B. Herndon, P.M. Gallagher, A.E. Barnato, J.P. Bynum, D.J. Gottlieb, E.S. Fisher, and J.S. Skinner. 2009. How much do patients' preferences contribute to resource use? *Health Affairs*, 28(3): 864-873.

Aron-Dine, A., L. Einav, and A. Finkelstein. 2013. The RAND health insurance experiment, three decades later. *Journal of Economic Perspectives*, 27(1): 197-222.

Aron-Dine, A., L. Einav, A. Finkelstein, and M. Cullen. 2015. Moral hazard in health insurance: do dynamic incentives matter? *Review of Economics and Statistics*, 97(4): 725-741.

Arrow, K.J. 1963. Uncertainties and the welfare economics of medical care *American Economic Review*, 53(5): 941-973.

Athey, S., and G.W. Imbens. 2017. The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives*, 31(2): 3-32.

Augurzky, B., T. Kopetsch, and H. Schmitz. 2013. What accounts for the regional differences in the utilisation of hospitals in Germany? *The European Journal of Health Economics*, 14(4): 615-627.

Baicker, K., and A. Chandra. 2004. Medicare spending, the physician workforce, and beneficiaries' quality of care. *Health Affairs*, 23(Suppl1): W4184-W4197.

Baicker, K., and D. Goldman. 2011. Patient cost-sharing and healthcare spending growth. *Journal of Economic Perspectives*, 25(2): 47-68.

Baicker, K., S.L. Taubman, H.L. Allen, M. Bernstein, J.H. Gruber, J.P. Newhouse, E.C. Schneider, B.J. Wright, A.M. Zaslavsky, and A.N. Finkelstein. 2013. The Oregon experiment—effects of Medicaid on clinical outcomes. *New England Journal of Medicine*, 368(18): 1713-1722.

Baker, L.C., M.K. Bundorf, and D.P. Kessler. 2014. Patients' preferences explain a small but significant share of regional variation in medicare spending. *Health Affairs* 33(6): 957-63.

Barreca, A.I., M. Guldi, J.M. Lindo, and G.R. Waddell. 2011. Saving babies? Revisiting the effect of very low birth weight classification. *Quarterly Journal of Economics*, 126(4): 2117-2123.

Bell, A., and K. Jones. 2015. Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data. *Political Science Research and Methods*, 3(1): 133-153.

Birkmeyer, J.D., B.N. Reames, P. McCulloch, A.J. Carr, W.B. Campbell, and J.E. Wennberg. 2013. Understanding of regional variation in the use of surgery. *The Lancet*, 382(9898): 1121-1129.

Bojke, C., A. Castelli, A. Street, P. Ward, and M. Laudicella. 2013. Regional variation in the productivity of the English national health service. *Health Econonomics*, 22(2): 194-211.

Brot-Goldberg, Z.C., A. Chandra, B.R. Handel, and J.T. Kolstad. 2017. What does a deductible do? The impact of cost-sharing on health care prices, quantities, and spending dynamics. *The Quarterly Journal of Economics*, 132(3): 1261-1318.

Brunner, B., and A. Kuhn. 2014. Announcement effects of health policy reforms: evidence from the abolition of Austria's baby bonus. *The European Journal of Health Economics*, 15(4): 373-388.

Callison, K., R. Kaestner, and J. Ward. 2020. Do supply-side forces explain geographic variation in health care use? *Economic Inquiry*, 59(1): 119-139.

Cantarero Prieto, D., and S. Lago-Penas. 2012. Decomposing the determinants of health care expenditure: the case of Spain. *The European Journal of Health Economics*, 13(1): 19-27.

Card, D., C. Dobkin, and N. Maestas. 2008. The impact of nearly universal insurance coverage on health care utilization: evidence from Medicare. *American Economic Review*, 98(5): 2242-58.

Card, D., C. Dobkin, and N. Maestas. 2009. Does Medicare save lives? *The Quarterly Journal of Economics*, 124(2): 597-636.

Card, D., D.S. Lee, Z. Pei, and A. Weber. 2015. Inference on causal effects in a generalized regression kink design. *Econometrica*, 83(6): 2453-2483.

Carpenter, C., and C. Dobkin. 2009. The effect of alcohol consumption on mortality: regression discontinuity evidence from the minimum drinking age. *American Economic Journal: Applied Economics*, 1(1): 164-182.

Chandra, A., D. Cutler, and Z. Song. 2011. Who ordered that? The economics of treatment choices in medical care. In *Handbook of Health Economics* Volume 2, edited by M.V. Pauly, T.G. McGuire and P.P. Barros, 397-432. Elsevier B.V.

Chandra, A., J. Gruber, and R. McKnight. 2010. Patient cost-sharing and hospitalization offsets in the elderly. *American Economic Review*, 100(1): 193-213.

Chandra, A., J. Gruber, and R. McKnight. 2014. The impact of patient cost-sharing on low-income populations: evidence from Massachusetts. *Journal of Health Economics*, 33: 57-66.

Chandra, A., and D.O. Staiger. 2007. Productivity spillovers in health care: evidence from the treatment of heart attacks. *Journal of Political Economy*, 115(1): 103-140.

Cockx, B., and C. Brasseur. 2003. The demand for physician services: evidence from a natural experiment. *Journal of Health Economics*, 22(6): 881-913.

Contoyannis, P., J. Hurley, P. Grootendorst, S.H. Jeon, and R. Tamblyn. 2005. Estimating the price elasticity of expenditure for prescription drugs in the presence of non-linear price schedules: an illustration from Quebec, Canada. *Health Economics*, 14(9): 909-923.

Corallo, A.N., R. Croxford, D.C. Goodman, E.L. Bryan, D. Srivastava, and T.A. Stukel. 2014. A systematic review of medical practice varition in OECD countries. *Health Policy*, 114: 5-14.

Cutler, D., J.S. Skinner, A.D. Stern, and D. Wennberg. 2019. Physician beliefs and patient preferences: a new look at regional variation in health care spending. *American Economic Journal: Economic Policy*, 11(1): 192-221.

Cutler, D.M., and R.J. Zeckhauser. 2000. The anatomy of health insurance. In *Handbook of Health Economics*, Volume 1, edited by A.J. Culyer and J.P. Newhouse, 563-643. Elsevier Science B.V.

Dalton, C.M., G. Gowrisankaran, and R.J. Town. 2020. Salience, myopia, and complex dynamic incentives: Evidence from Medicare Part D. *The Review of Economic Studies*, 87(2): 822-869.

Dartmouth Atlas Project. 2020. Dartmouth Atlas Data [Online database]. The Dartmouth Institute for Health Policy and Clinical Practice. https://atlasdata.dartmouth.edu accessed Sept 2020.

Doyle Jr, J.J., J.A. Graves, J. Gruber, and S.A. Kleiner. 2015. Measuring returns to hospital care: Evidence from ambulance referral patterns. *Journal of Political Economy*, 123(1): 170-214.

Einav, L., and A. Finkelstein. 2018. Moral hazard in health insurance: What we know and how we know it. *Journal of the European Economic Association*, 16(4): 957-982.

Einav, L., A. Finkelstein, and P. Schrimpf. 2015. The response of drug expenditure to nonlinear contract design: Evidence from Medicare Part D. *The Quarterly Journal of Economics*, 130(2): 841-899.

Eliason, M., P. Johansson, and M. Nilsson. 2019. Forward-looking moral hazard in social insurance. *Labour Economics*, 60: 84-98.

Eurostat European Commission. 2018. Regions in the European Union - Nomenclature of territorial units for statistics NUTS 2016. *Eurostat Methodologies & Working papers*. Luxembourg: Publications offices of the European Union.

Farbmacher, H., and J. Winter. 2013. Per-period co-payments and the demand for health care: evidence from survey and claims data. *Health Economics*, 22(9): 1111-1123.

Feng, J., H. Song, and Z. Wang. 2020. The elderly's response to a patient cost-sharing policy in health insurance: Evidence from China. *Journal of Economic Behavior & Organization*, 169: 189-207.

Filippini, M., G. Masiero, and K. Moschetti. 2006. Socioeconomic determinants of regional differences in outpatient antibiotic consumption: evidence from Switzerland. *Health Policy*, 78(1): 77-92.

Finkelstein, A., M. Gentzkow, and H. Williams. 2016. Sources of geographical variation in health care: Evidence from patient migration. *The Quarterly Journal of Economics*: 1681-1726.

Finkelstein, A., S. Taubman, B. Wright, M. Bernstein, J. Gruber, J.P. Newhouse, H. Allen, and K. Baicker. 2012. The Oregon health insurance experiment: evidence from the first year. *The Quarterly Journal of Economics*, 127(3): 1057-1106.

Fisher, E., D. Goodman, J. Skinner, and K. Bronner. 2009. Health care spending, quality, and outcomes: more isn't always better. *Dartmouth Atlas Project Topic Brief.* Hanover, NH: The Dartmouth Institute for Health Policy and Clinical Practice.

Fisher, E.S., D.E. Wennberg, T.A. Stukel, D.J. Gottlieb, F.L. Lucas, and E.L. Pinder. 2003. The implications of regional variations in Medicare spending. Part 2: health outcomes and satisfaction with care. *Annals of Internal Medicine*, 138(4): 288-298.

Fukushima, K., S. Mizuoka, S. Yamamoto, and T. Iizuka. 2016. Patient cost sharing and medical expenditures for the Elderly. *Journal of Health Economics*, 45: 115-130.

Glover, J.A. 1938. The incidence of tonsillectomy in school children. *Proceedings of the Royal Society of Medicine*, 31(10): 1219-36.

Godøy, A., and I. Huitfeldt. 2020. Regional variation in health care utilization and mortality. *Journal of Health Economics*, 71: 102254.

Gottlieb, D.J., W. Zhou, Y. Song, K.G. Andrews, J.S. Skinner, and J.M. Sutherland. 2010. Prices don't drive regional Medicare spending variations. *Health Affairs*, 29(3): 537-543.

Guo, A., and J. Zhang. 2019. What to expect when you are expecting: Are health care consumers forward-looking? *Journal of Health Economics*, 67: 102216.

Gusmano, M.K., D. Weisz, V.G. Rodwin, J. Lang, M. Qian, A. Bocquier, V. Moysan, and P. Verger. 2014. Disparities in access to health care in three French regions. *Health Policy*, 114(1): 31-40.

Göpffarth, D., T. Kopetsch, and H. Schmitz. 2016. Determinants of regional variation in health expenditures in Germany. *Health Economics*, 25(7): 801-15.

Han, H.-W., H.M. Lien, and T.-T. Yang. 2019. Patient cost sharing and healthcare utilization in early childhood: evidence from a regression discontinuity design. *American Economic Journal: Economic Policy*, 12(3): 238-278.

Hayen, A., T.J. Klein, and M. Salm. 2018. Does the framing of patient cost-sharing incentives matter? The effects of deductibles vs. no-claim refunds. *IZA Institute of Labor Economics Discussion Paper*, IZA DP No. 11508.

Heijink, R., P. Engelfriet, C. Rehnberg, S.A. Kittelsen, U. Hakkinen, and H.s.g. Euro. 2015. A window on geographic variation in health care: insights from EuroHOPE. *Health Economics*, 24 Suppl 2: 164-77.

Henricson, K., E. Melander, S. Mölstad, J. Ranstam, B. Hanson, G. Rametsteiner, P. Stenberg, and A. Melander. 1998. Intra-urban variation of antibiotic utilization in children: influence of socio-economic factors. *European Journal of Clinical Pharmacology*, 54(8): 653-657.

Hovstadius, B., B. Åstrand, and G. Petersson. 2010. Assessment of regional variation in polypharmacy. *Pharmacoepidemiology and Drug Safety*, 19(4): 375-383.

Jakobsson, N., and M. Svensson. 2016a. Copayments and physicians visits: A panel data study of Swedish regions 2003-2012. *Health Policy*, 120(9): 1095-1099.

Jakobsson, N., and M. Svensson. 2016b. The effect of copayments on primary care utilization: results from a quasi-experiment. *Applied Economics*, 48(39): 3752-3762.

Johansson, N., N. Jakobsson, and M. Svensson. 2018. Regional variation in health care utilization in Sweden–the importance of demand-side factors. *BMC Health Services Research*, 18: 403.

Johansson, N., N. Jakobsson, and M. Svensson. 2019. Effects of primary care cost-sharing among young adults: varying impact across income groups and gender. *The European Journal of Health Economics*, 20(8): 1271-1280.

Johnson, W.C., and J.F. Biniek. 2020. Sources of geographic variation in health care spending among individuals with employer sponsored insurance. *Medical Care Research and Review*: 1077558720926095.

Keeler, E.B., and J.E. Rolph. 1988. The demand for episodes of treatment in the health insurance experiment. *Journal of Health Economics*, 7(4): 337-367.

Klein, T.J., M. Salm, and S. Upadhyay. 2020. The reponse to dynamic incentives in insurance contracts with a deductible: evidence from a differences-in-regression-discontinuities design. *IZA Institute of Labor Economics Discussion Paper*, IZA DP No. 13108.

Kopetsch, T., and H. Schmitz. 2014. Regional variation in the utilisation of ambulatory services in Germany. *Health Economics*, 23(12): 1481-92.

Kozyrskyj, A.L. 2002. Prescription medications in Manitoba children. *Canadian Journal of Public Health*, 93(2): S63-S69.

Kunz, J.S., and R. Winkelmann. 2017. An econometric model of healthcare demand with nonlinear pricing. *Health Economics*, 26(6): 691-702.

Lavergne, M.R., M. Barer, M.R. Law, S.T. Wong, S. Peterson, and K. McGrail. 2016. Examining regional variation in health care spending in British Columbia, Canada. *Health Policy*, 120(7): 739-748.

Lee, D.S., and T. Lemieux. 2010. Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, 48(2): 281-355.

Lemieux, T., and K. Milligan. 2008. Incentive effects of social assistance: A regression discontinuity approach. *Journal of Econometrics*, 142(2): 807-828.

Manning, W.G., J.P. Newhouse, N. Duan, E.B. Keeler, and A. Leibowitz. 1987. Health insurance and the demand for medical care: evidence from a randomized experiment. *American Economic Review*, 77(3): 251-277.

Molitor, D. 2018. The evolution of physician practice styles: evidence from cardiologist migration. *American Economic Journal: Economic Policy*, 10(1): 326-56.

Moura, A., M. Salm, R. Douven, and M. Remmerswaal. 2019. Causes of regional variation in Dutch healthcare expenditures: Evidence from movers. *Health Economics*, 28(9): 1088-1098.

Newhouse, J.P. 2014. Introduction. In A. Finkelstein. *Moral hazard in health insurance*. New York: Columbia University Press.

Nilsson, A., and A. Paul. 2018. Patient cost-sharing, socioeconomic status, and children's health care utilization. *Journal of Health Economics*, 59: 109-124.

OECD. 2014. Geographic variations in health care: What do we know and what can be done to improve health system preformance? *OECD Health policy studies*. Organisation for Economic Co-operation and Development.

OECD. 2019. *Health at a glance 2019: OECD indicators*. Paris: OECD Publishing.

OECD. 2020. OECD Health Data 2020. [Online database]. Organisation for Economic Co-operation and Development. http://stats.oecd.org/Index.aspx?DataSetCode=SHA accessed Sept 2020.

Olsen, C.B., and H.O. Melberg. 2018. Did adolescents in Norway respond to the elimination of copayments for general practitioner services? *Health Economics*, 27(7): 1120-1130.

Osika Friberg, I., G. Krantz, S. Määttä, and K. Järbrink. 2016. Sex differences in health care consumption in Sweden: a register-based cross-sectional study. *Scandinavian Journal of Public Health*, 44(3): 264-273.

Otto, M., P. Armeni, and C. Jommi. 2018. Variations in non-prescription drug consumption and expenditure: determinants and policy implications. *Health Policy*, 122(6): 614-620.

Pauly, M.V. 1968. The economics of moral hazard: comment. *American Economic Review*, 58(3): 531-537.

Reich, O., C. Weins, C. Schusterschitz, and M. Thoni. 2012. Exploring the disparities of regional health care expenditures in Switzerland: some empirical evidence. *European Journal of Health Economics*, 13(2): 193-202.

Remmerswaal, M., J. Boone, M. Bijlsma, and R. Douven. 2019a. Cost-sharing design matters: A comparison of the rebate and deductible in healthcare. *Journal of Public Economics*, 170: 83-97.

Remmerswaal, M., J. Boone, and R. Douven. 2019b. Selection and moral hazard effects in healthcare. *CPB Discussion Paper*: Netherlands Bureau for Economic Policy Analysis.

Reschovsky, J.D., J. Hadley, and P.S. Romano. 2013. Geographic variation in fee-for-service medicare beneficiaries' medical costs is largely explained by disease burden. *Medical Care Research and Review*, 70(5): 542-563.

RKA. 2020. Kolada Swedish municipality and regional database [Online database]. Rådet för främjande av kommunala analyser. www.kolada.se accessed Sept 2020.

Salm, M., and A. Wübker. 2020. Sources of regional variation in healthcare utilization in Germany. *Journal of Health Economics*, 69: 102271.

Schleiniger, R. 2014. Health care cost in Switzerland: quantity- or price-driven? *Health Policy*, 117(1): 83-9.

Schreyögg, J., and M.M. Grabka. 2010. Copayments for ambulatory care in Germany: a natural experiment using a difference-in-difference approach. *The European Journal of Health Economics*, 11(3): 331-341.

SFS. 2008:962. *Lag om valfrihetssystem*. Finansdepartementet

SFS. 2017:30. *Hälso- och sjukvårdslag*. Socialdepartementet.

Sheiner, L. 2014. Why geographic variation in health care spending cannot tell us much about the efficiency or quality of our health care system. *Brookings paper on economic activity*, Fall 2014.

Shigeoka, H. 2014. The effect of patient cost sharing on utilization, health, and risk protection. *American Economic Review*, 104(7): 2152-2184.

Simonsen, M., L. Skipper, and N. Skipper. 2016. Price sensitivity of demand for prescription drugs: exploiting a regression kink design. *Journal of Applied Econometrics*, 31(2): 320-337.

Skinner, J. 2011. Causes and consequenses of regional variations in health care. In *Handbook of Health Economics*, Volume 2, edited by Mark V. Pauly, P.P. Barros and T.G. Mcguire, 45-93. London: Elsevier Science.

SKR. 2020. *Patientavgifter i öppen hälso- och sjukvård år 2020*. Swedish Association of Local Authorities and Regions. https://skr.se/download/18.1509f18f17059700521f0222/1582650823232/avgifter_%C3%B6ppen_slutenvard_2020_uppdaterad.pdf accessed Sept 2020.

Song, Y., J. Skinner, J. Bynum, J. Sutherland, J.E. Wennberg, and E.S. Fisher. 2010. Regional variations in diagnostic practices. *New England Journal of Medicine*, 363(1): 45-53.

Sutherland, J.M., E.S. Fisher, and J.S. Skinner. 2009. Getting past denial--the high cost of health care in the United States. *New England Journal of Medicine*, 361(13): 1227.

TLV. 2020. Högkostnadsskyddet. Tandvårds- och läkemedelsförmånsverket. https://www.tlv.se/lakemedel/hogkostnadsskyddet.html accessed Sept 2020.

Wennberg, J., and A. Gittelsohn. 1973. Small area variations in health care delivery: a population-based health information system can guide planning and regulatory decision-making. *Science*, 182(4117): 1102-1108.

Winkelmann, R. 2004. Co-payments for prescription drugs and the demand for doctor visits–Evidence from a natural experiment. *Health Economics*, 13(11): 1081-1089.

Wooldridge, J.M. 2014. *Introduction to econometrics*. Hampshire: Cengage Learning EMAE.

Zhang, Y., K. Baicker, and J.P. Newhouse. 2010a. Geographic variation in Medicare drug spending. *New England Journal of Medicine*, 363(5): 405-409.

Zhang, Y., K. Baicker, and J.P. Newhouse. 2010b. Geographic variation in the quality of prescribing. *New England Journal of Medicine*, 363(21): 1985.

Zhang, Y., S.H. Baik, A.M. Fendrick, and K. Baicker. 2012. Comparing local and regional variation in health care spending. *New England Journal of Medicine*, 367(18): 1724-1731.

Ziebarth, N.R. 2010. Estimating price elasticities of convalescent care programmes. *The Economic Journal*, 120(545): 816-844.

Zuckerman, S., T. Waidmann, R. Berenson, and J. Hadley. 2010. Clarifying sources of geographic differences in Medicare spending. *New England Journal of Medicine*, 363(1): 54-62.

Zweifel, P., and W.G. Manning. 2000. Moral hazard and consumer incentives in health care. In *Handbook of Health Economics*, Volume 1, edited by A.J. Culyer and J.P. Newhouse, 409-459. Elsevier Science B.V.