Health, Insurance and Expenditures. Four Essays in Empirical Health Economics.

Dissertation

zur Erlangung des akademischen Grades eines
Doktors der Wirtschaftswissenschaften
(Dr. rer. pol.)
durch die Fakultät für Wirtschaftswissenschaften der
Universität Duisburg-Essen
Campus Essen

vorgelegt von
Name: Florian Klohn
Ort: Iserlohn
Essen, 2015

Tag der mündlichen Prüfung: 07.09.2016
Erstgutachter: Prof. Martin Karlsson, Ph.D.
Zweitgutachter: Prof. Dr. Jürgen Wasem

Acknowledgements

I would like to thank my supervisor, co-author and former employer Martin Karlsson for his inspiration and his valuable research support. In addition, I am grateful to my second supervisor, Jürgen Wasem, for reviewing this thesis and Ben Rickayzen for his contribution to our joint paper. Special thanks go to my friend and co-author Christoph Strupat for his enriching discussions on topics both related and unrelated to our joint research.

I would also like to thank research fellows and former colleagues such as Stefan Pichler, Stefan Goldbach, Maike Schmitt, Hendrik Schmitz, Claudia Andreella, Sarah Okoampah, Simon Decker, Martin Fischer, Maksim Obrizan and Nana Chanishvili-Schön for their fruitful and supportive discussions during my time at the University of Duisburg-Essen and TU-Darmstadt. I also like to acknowledge the interesting and helpful ideas which I gathered at scientific conferences, workshops and from anonymous referees on some of my research.

Most importantly, I would like to thank my parents and my sister for their love and for supporting me in countless ways throughout my life.

Contents

List of Figures

List of Tables

Chapter 1

Motivation and Contribution

1.1 Background

Individual health directly reflects fundamental physical or psychological needs and is very important from an economic perspective. For example, health is positively correlated with important areas of life such as social participation (e.g. Snelgrove et al., 2009) or membership of the workforce (e.g. García-Gómez et al., 2010). The importance of health has been recognised by the WHO, which defines it as 'a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity'[1]. The relevance of disease is also recognised at the population level, particularly within the public health sphere, but also because of the impact on public spending (e.g. Getzen, 2000a) and economic growth (e.g. Karlsson et al., 2014).

The financial importance of health can be seen at the aggregated level by considering the amount of money spent on health and health-related initiatives. Figure 1.1 highlights the increasing economic importance of the healthcare sector from an international perspective. Although there is heterogeneity within the WHO health regions[2], healthcare expenditure as a share of GDP increased for all regions during the observed period.

[1] Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 19-22 June, 1946; signed on 22 July 1946 by the representatives of 61 States (Official Records of the World Health Organization, no. 2, p. 100) and entered into force on 7 April 1948.

[2] Australia and China, for example, are included in the Western Pacific region.

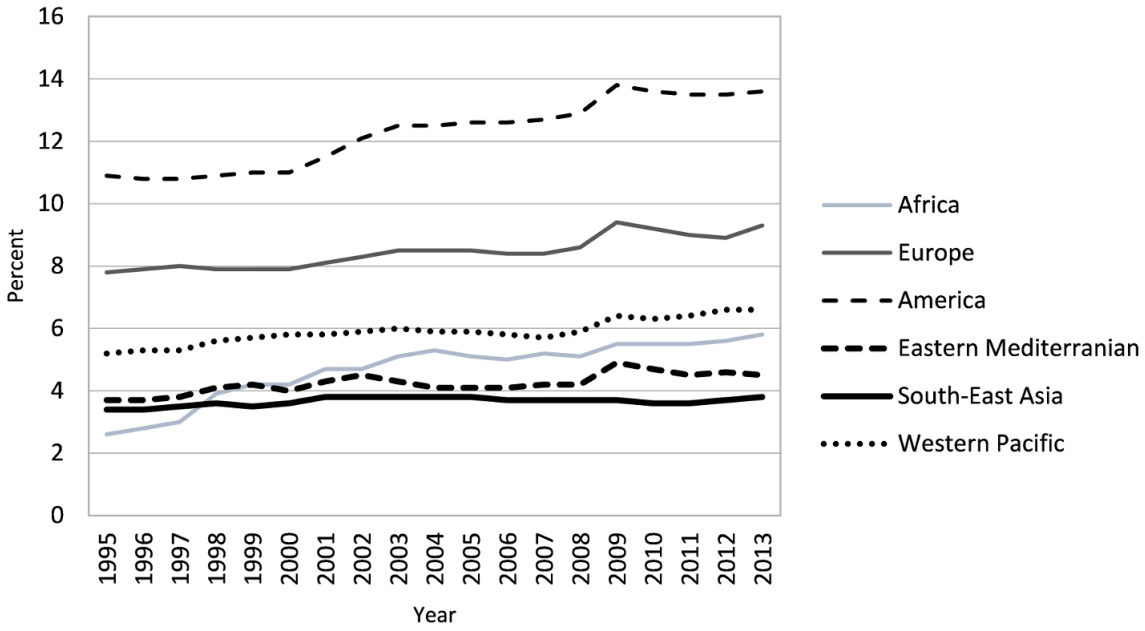


Figure 1.1 Total health expenditure as a percentage of GDP (Data source: WHO)

However, Southeast Asia has seen only a very modest increase, and the Eastern Mediterranean region displayed a small decrease from 2009 onwards. The strongest increase is seen in Africa, which may be explained to a certain extent by very low baseline healthcare expenditures, the adoption of professional medical procedures during the observed period, and increased financing from international development programs that emphasise health as a major driver of economic development (Ravishankar et al., 2009). Overall, there is a pattern of increased relative healthcare expenditure over the observed period, although different mechanisms explain variations within and between regions. Comparing regions highlights the long-recognised pattern that the relative quantity of healthcare expenditure is greatest in the most economically developed countries (e.g. Farag et al., 2012), with America at the top and Southeast Asia at the bottom (from 1998 onwards). High base levels and a subsequent increase in healthcare expenditure in America are mainly explained by a mixture of technological progress and a market-oriented approach to the structure of their healthcare sector (Chandra and Skinner, 2012).

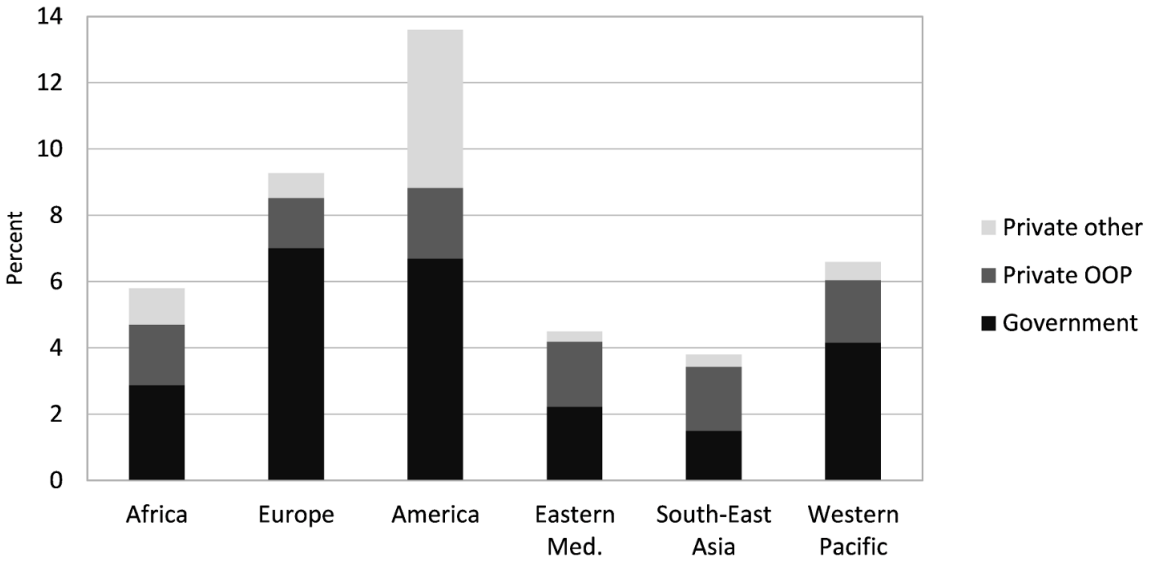


Figure 1.2 Expenditures on health as a percentage of gross domestic product (Year: 2013, Data source: WHO)

Turning to the sources used to finance healthcare expenditures in our six geographical regions, Figure 1.2 distinguishes between government expenditures and private expenditures in 2013. Private expenditures are represented by out-of-pocket (OOP) payments and other private sources such as private insurance. Overall, we find that the role of government is emphasised in Europe, where private sources play only a minor role. America has a relatively large share of private healthcare expenditures based on private insurance markets, while in less-developed regions (Africa, Eastern Mediterranean, Southeast Asia) OOP payments play a large role. OOP payments also include financing via private transfer and family networks. As recognised by Esping-Andersen (1990) in his classification of welfare systems, families generally play a very important role in providing healthcare and financing social security[3]. OOP payments are less important in Europe and America relative to the other regions.

Overall, European healthcare expenditures are much more reliant on government (including social insurance) than in America and other areas (the latter also have much lower levels of health expenditures overall). Although the sources of healthcare expenditures are quite different across the six regions, on average the relative

[3]Esping-Andersen (1990) classifies the welfare state based on whether social security is provided mainly by the market, the government or reciprocity (including social insurance and family networks).

importance of healthcare expenditures, given financial ability, has increased across every region over the last 20 years. Increasing demand for healthcare, the arrival of new technologies and an increase in unhealthy lifestyles are a burden for fiscal budgets, as healthcare expenditures have grown strongly in recent decades in both developed and less-developed regions. If this trend continues, it may cause problems for future decision making because individuals, companies, and politicians have to allocate resources under financial restrictions.

1.2 Contribution and Agenda

Knowledge of how to control and properly predict healthcare expenditures is essential for good policy advice. This dissertation employs empirical analysis using data from different countries to provide multiple contributions to the contemporary literature on managing and projecting healthcare expenditures, which are very important for both developed and developing countries. The dissertation is organised into five chapters. Chapter 1 contains this introduction. Chapter 2 empirically emphasises the importance of ageing and morbidity for future long-term care (LTC) expenditures. Chapter 3 discusses the role of individual characteristics that may serve as barriers to participating in public health programmes which may be used to reduce healthcare expenditures. Chapter 4 contributes to the empirical detection of asymmetric information (IA) in private insurance markets, which are widely acknowledged to induce inefficient resource allocation. Finally, Chapter 5 examines from a development perspective whether introducing a formal, nationwide health insurance scheme affects participation in informal transfer mechanisms. The following provides a more detailed description of our research agenda in the context of contemporary literature.

Chapter 2: Ageing, Time-To-Death and Care Costs for Older People in Sweden

Chapter 2 addresses the ongoing scientific discussion surrounding evidence-based planning of future healthcare expenditures and the impact of ageing on healthcare expenditures. This is a particular issue in relatively developed countries, which are generally seeing a decline in fertility combined with increasing life expectancies (e.g. Herzer et al., 2012). We add to the health economics literature since Zweifel et al. (1999), investigating whether age itself is still an important predictor of healthcare expenditures, once time-to-death (TTD) is accounted via empirical analysis. There

are various methodological problems with estimating the impact of an ageing population on future budgets, and few studies are based on aggregated data. This field of research is also important in the context of LTC expenditures (e.g. Larsson et al., 2008). To shed new light on this issue, we use aggregated data from official statistics in Sweden, a country where social care is largely financed by taxes, to estimate the impact of age on LTC expenditures. Our empirical analysis applies panel data methods and accounts for end-of-life morbidity, which allows us to separate age and TTD effects to derive projections of expected future LTC expenditures. We also allow for differences between domiciliary and institutional LTC in our analysis, which may prove important if different treatment paths have a heterogeneous impact on resource allocation or healthcare provision. A proper understanding of the association between age and (non curative) healthcare expenditures is key to adequately projecting the financial consequences of population ageing on future healthcare budgets. The main contribution of this chapter is that we provide a theoretically more convincing measure for end-of-life morbidity based on population level mortality than contemporary literature does.

Chapter 3: The Ability to Memorise and Participation in the English Bowel Cancer Screening Programme

The third chapter emphasises how policies designed to reduce healthcare expenditures need to incorporate the characteristics of their specific target group to be successful. It is widely acknowledged that societies where individuals do not have to (fully) pay a fair premium must finance health expenditures at a societal level (Mullainathan et al., 2012). Healthcare programmes may be able to positively influence healthcare decisions (from a policy perspective). However, in libertarian societies the success of such programmes depends on each individual's decision to participate or not. Hence, the individual characteristics of a specific target group which affect participation in such programmes must be carefully identified and accounted for when designing policies. For example, in programmes designed for the elderly, a decline in physical and cognitive abilities may be a barrier to participation. In this context, we use individual-level data and apply regression and matching methods to empirically assess whether an individual's declining ability to memorise information decreases their likelihood of participating in a nationwide healthcare programme for bowel cancer screening in England. The main contribution of this chapter is further emphasis that public health programmes need to be designed with careful regard to the specifics of their target group. We also contribute to the psychological

and behavioural economics literature that assesses unexpected health-related decision making processes (e.g. Mullainathan et al., 2012; Chetty, 2015). Our findings may help further develop and enhance models to better explain real-world health economic decision making.

Chapter 4: Heterogeneous Parameters and Detecting Selection Based on 'Unused Characteristics' in Private Health Insurance Markets

The fourth chapter investigates the empirical detection of selection in private insurance markets[4]. Detecting selection within insurance markets is an important regulatory issue, because it is often driven by asymmetric information, and limited resources make it economically essential to find ways to efficiently provide and finance healthcare. From a health economics perspective, it is widely acknowledged that health policies must be both efficient and equitable (Culyer and Wagstaff, 1993). This can be seen, for example, in the UK, where healthcare financing and provision is built to a large extent on the National Health Service (NHS), with optional additional private health insurance to incorporate the advantages of private markets into the system. Additional private health insurance can be helpful if it allows for specific services or treatments which are not necessary from a health policy perspective but which match an individual's specific needs and may help satisfy patients and enhance technological progress, even if public or social insurance guarantees a relatively high standard of healthcare provision (Colombo and Tapay, 2004).

For private insurance markets, the literature on selection and inefficiencies has gained a lot of attention in the economics literature, following Akerlof (1970) and Rothschild and Stiglitz (1976). Further issues surrounding information asymmetries and selection in insurance markets, such as moral hazard, cream skimming or propitious selection have also been considered (e.g. Leidl, 2008; Einav and Finkelstein, 2011). Despite numerous theoretical discussions about selection mechanisms in private insurance markets that can imply inefficiencies, the empirical detection of selection based on expected risk is not yet conclusive. The main contribution of this chapter is a discussion of the empirical detection of information asymmetries (IAs) based on approaches proposed in Finkelstein and McGarry (2006). We formally and empirically show that indirect assessment of IAs using 'unobserved' (Finkelstein and

[4]There is also a wide-ranging debate about the role of risk selection in public health insurance systems (e.g. Nuscheler and Knaus, 2005; van de Ven et al., 2007; Bauhoff, 2012), which is crucial for designing effective risk adjustment schemes. Although we regard the technical issues surrounding selection to be potentially relevant to such frameworks as well, we only discuss our ideas in the context of private health insurance, following the literature we respond to.

McGarry, 2006) or 'unused' variables (Finkelstein and Poterba, 2014) can lead to faulty conclusions about the direction of selection being drawn. We criticise standard approaches used to detect IAs based on 'unused characteristics' and suggest a remedy for their shortcomings. Specifically, we show that the approach suggested by Finkelstein and McGarry (2006) is problematic if the parameters of interest are driven by heterogeneous sub-populations with different outcomes for an 'unobserved characteristic'. We empirically showcase our idea by using simulated data and provide an empirical application about selection in a private health insurance market with survey data from England. Our findings are of major importance for designing and evaluating regulatory frameworks in private (health) insurance markets where selection is supposed to be based on specific characteristics.

Chapter 5: The Relationship Between Public Health Insurance and Informal Transfer Networks in Ghana

Chapter five discusses the relationship between formal insurance and informal transfer networks. Private healthcare expenditures based on out-of-pocket (OOP) payments are a common practice across the world. In private insurance markets, the contract design usually implements partial financing via OOP payments to reduce the well-known problem of moral hazard (Cutler and Zeckhauser, 2000). However, in the absence of public or private health insurance, OOP payments play the most important role in healthcare expenditures. In this context, leaving aside any concerns about equity, healthcare financing with a small risk pool may be inefficient due to high uncertainty and risk costs (e.g. Martinez-Giralt and Barros, 2013). Recently, several developing countries have introduced public health insurance to reduce shortcomings in such informal insurance networks and improve healthcare provision (Wagstaff, 2010). Although public and social health insurance are playing an increasing role in developing countries, private health expenditures based on OOP payments are still the dominant way to pay for healthcare (Dye et al., 2013). In the absence of formal institutions that provide risk sharing, informal networks can naturally evolve and provide remedy. In such a setting, the role of informal help from the community, especially family, is very important. However, this kind of risk sharing is also criticised in the development economics literature as it can constrain economic growth (e.g. Grimm et al., 2013).

We contribute to the health insurance literature in a development context (e.g. De Weerdt and Fafchamps, 2011; Powell-Jackson et al., 2014), with a focus on the relationship between formal and informal health insurance markets (e.g. Landmann

et al., 2012; Lin et al., 2014). We empirically assess whether implementing the national health insurance scheme in Ghana in 2005 crowded out informal transfer networks in the short run. In addition, we assess whether the new insurance scheme contributes to health outcomes and individual healthcare expenditures. The main contribution of this chapter is our identification strategy, which uses the quasi-exogenous variation in implementing the programme at the regional level. The relationship between different institutions/markets is also important for designing successful policies in more developed countries. Hence, our research question is also of interest for health economic policies beyond the developing world.

The next section outlines the methods and findings for each of the main chapters.

1.3 Methods and Findings

Chapter 2: Ageing, Time-to-Death and Care Costs for Older People in Sweden

We test the 'red herring' hypothesis in the context of Swedish LTC expenditures by using municipality-level panel data based on administrative records for 1998 to 2008. We restrict our sample to all municipalities that exclusively provided LTC services to allow us to focus solely on non-curative healthcare expenditures. Our sample covers half of the Swedish population. We investigate whether end-of-life morbidity (or TTD) is a better predictor of LTC expenditures than age by controlling for contemporary and future mortality rates. We derive a retrospective construct of TTD that captures mortality within two years on the aggregated level and allows us to account for individual end-of-life morbidity by using future mortality rates with respect to the population of interest in a given year. We apply a fixed effects estimator to control for potential endogeneity due to unobserved heterogeneity between municipalities.

Overall, we find ageing to be the most important driver of LTC expenditures in our empirical specifications. In sub-sample estimates, we further distinguish between institutional and domiciliary LTC provision and also provide sex- and age-specific estimates. The resulting findings suggest that individuals switch between domiciliary and institutional care at the end of their lives. Sex-specific specifications also show a negative relationship between end-of-life morbidity and domiciliary LTC expenditures, mostly driven by women. The latter finding may indicate that informal care also plays an important role in the provision and financing of LTC expenditures in Sweden. Age-specific TTD effects show that the impact of TTD on LTC expenditures

can mostly be explained by a relatively young cohort, the 70- to 74-year-olds. Overall, Swedish age structure remains a very important driver of total LTC payments. In the context of Sweden, the high relevance of ageing for financial budgets is also emphasised by our projections for LTC expenditures, which we calculate based on our estimates and by 'increasing' life expectancy by one year. Based on our findings and predictions of Swedish age structure from Statistics Sweden, we also calculated cost projections for a period of 100 years. From an economic perspective, these costs are considerable.

Chapter 3: The Ability to Memorise and Participation in the English Bowel Cancer Screening Programme

To analyse the causal impact of the ability to memorise information on participation in the national bowel cancer screening programme in England, we use panel data from the ELSA survey, which is representative of the elderly population. We use an objective measure for an individual's decline in memory and apply linear regression and propensity score matching in our identification strategy to estimate the impact of a decline in memorisation ability on participating in bowel cancer screening during the years following the artificial treatment. We apply covariate adjustment by controlling for important confounders and account for the timings of when the data was generated to mitigate problems such as omitted variable bias and reverse causality. We also apply a doubly robust regression approach, which is widely used in the epidemiological literature and has recently been adopted in the empirical health economics literature (e.g. Schmitz and Westphal, 2015). This provides unbiased estimates when either the treatment or outcome equation of our empirical model is correctly specified. Dealing with endogeneity is important in our context, since our treatment – reduced ability to memorise – is probably not randomised during the survey period. We provide evidence of a negative relationship between our treatment and participation in bowel cancer screening. Our analysis also distinguishes between different domains of memorisation and finds that the main effect can be largely explained by the ability to remember information with some delay. Our findings are primarily of interest for health policy design and in the context of medical decision making. If a public health programme is designed to increase health or reduce expenditures, the particulars of its target group should be considered to efficiently achieve the programme goals.

Chapter 4: Heterogeneous Parameters and Detecting Selection Based on 'Unused Characteristics' in Private Health Insurance Markets

The literature following Finkelstein and McGarry's (2006) 'unused characteristics' approach is based on the idea that there are variables which insurance companies do not use to calculate risk premiums, but which can be used in an empirical research analysis. It argues that if these variables hold explanatory power for both an individual's risk and insurance status, the signs of the estimated coefficients can be used to indirectly detect information asymmetries (IAs). These approaches are important when testing whether a specific variable introduces selection into insurance markets. We formally and empirically show that these standard unused characteristics approaches (e.g. Finkelstein and McGarry, 2006) can be problematic if the resulting evidence is wrongly based on mean coefficients. We highlight this issue by allowing for heterogeneity in individual parameters. In doing so, we show that a negative correlation between the coefficients of interest may lead to faulty conclusions if adverse selection was detected in the first place, while a positive correlation may lead to a faulty conclusion about advantageous selection into an insurance company's risk pool. Obviously, the unused characteristics approach can also lead to faulty conclusions if no selection is detected. We bolster the relevance of this finding by simulating different correlation structures between a hypothetical unused characteristic and both insurance and risk status, allowing for individual heterogeneity in the data generating process. Using a multilevel model for our estimates, we find that standard unused characteristics approaches do not reveal this kind of heterogeneity under the circumstances discussed. We also provide an empirical implementation for the (voluntary) private health insurance market in England. We use panel data from between 2002 and 2013 from the ELSA survey and choose an individual's available time as an unused characteristic, as this variable is clearly not available to insurance companies for pricing purposes. Our findings show that individual parameter heterogeneity is relevant to real markets as well. Although adverse selection into the insurance market is empirically detected, this adverse selection should be interpreted with caution, since the estimated parameters are negatively correlated. Our findings show that although contemporary literature (Einav et al., 2009) focuses on calculating welfare effects in insurance markets due to IAs, the empirical identification of IAs is still an important research area. We conclude that knowledge about the empirical detection of selection in insurance markets is inconclusive, and that future empirical analysis should account for the correlation structures of coefficients when applying an unused characteristics approach.

Chapter 5: The Relationship Between Public Health Insurance and Informal Transfer Networks in Ghana

To assess the influence of the new Ghanaian health insurance scheme (NHIS), we analyse cross-sectional data from the fifth Ghanaian Living Standard Survey, which is representative for the 2005 Ghanaian population. As the NHIS was implemented at the district level, we can use the variation in when individuals were interviewed at the various sub-districts. Since most districts introduced the NHIS during the survey period in 2005, we compare individuals who were interviewed before and after the introduction of the insurance scheme, conditional on fixed effects of district and interview month. Using the quasi-exogenous variation in the availability of formal health insurance, we apply ordinary least squares estimates to estimate the impact of the new insurance scheme on making and receiving transfers at the extensive and intensive margin. We also test whether the treatment effect depends on the relationship between the recipient and donor of a transfer. Our empirical findings indicate that introducing a formal health insurance scheme reduces the probability of making transfers. In addition, the number of remittances decreases significantly. We also find that a relatively close relationship between the recipient and donor of a transfer, holding everything else constant, reduces the crowding out. One potential explanation for this heterogeneity is sharing obligations, which are known to be strong in developing countries and are recognised by the contemporary literature as being a barrier to economic growth (e.g. Grimm et al., 2013). The decrease in informal transfer networks due to formal insurance is found to be lowest in kinship networks; therefore, our findings raise the question of whether formal insurance can overcome this issue, at least in the short run. Our analysis of health-related outcomes suggests that the NHIS reduces respondents' OOP expenditures, which is in line with our expectations. Overall, our findings indicate that public health insurance schemes strongly affect how healthcare services are paid for and may also support economic development in the long run. However, the results also emphasise the effects of new policies on existing institutions, which can be an important issue in many health policy contexts.

Chapter 2

Ageing, Time-To-Death and Care Costs for Older People in Sweden[1]

2.1 Introduction

Most developed countries have ageing populations, which has implications for public spending on long-term care (LTC) and healthcare. Recent population projections for Sweden suggest an increase in the old-age dependency ratio (population aged 65+ per 100 persons aged 15-64) from 32 in 2015 to 39 in 2050 (United Nations, 2013), placing an enormous burden on public spending.

An ageing society is expected to lead to an increased demand for care services, raising concerns about the sustainability of financing these services. Indeed, there may be no magic wand that guarantees both the availability and quality of LTC provision without continuous cost increases (Meier and Werding, 2010). Precise measurement of the influence of ageing on care costs is essential to make reliable projections of demand for care services. An oft-discussed issue when quantifying the impact of ageing on public funding is the possibility of omitted variables. If the period of increased care needs is simply postponed as life expectancy increases, age itself has limited explanatory power in projecting care expenditures. To circumvent this problem empirically, older people's time-to-death (TTD) can be considered. This study contributes to the literature in four different ways: First, we provide a new TTD measure to estimate its impact on total, institutional and domiciliary LTC costs, using high-quality administrative data aggregated at the municipality level. Unlike existing studies, which are implemented at the macro level and only use raw

[1]This study is a joint work with Martin Karlsson. See Karlsson and Klohn (2014) for the corresponding published paper.

mortality rates to measure TTD, our approach accounts for the probability of dying within a specific timeframe. Using macro-level data means the findings are representative for the population as a whole. Second, we test the extent to which costs due to TTD vary between different age groups and identify heterogeneity between men and women. Third, we combine our estimates with freely available population projections from Statistics Sweden to calculate the future financial consequences of the expected demographic change. Sweden provides us with an ideal institutional background for our analysis: As Sweden balances funding on a national level based on needs, we can expect regional funding to solely reflect local needs. Hence, and fourth, our explanatory variables do not reflect differences in budget restrictions, a common problem with macro-level analysis.

This chapter is organised as follows: In the next section, we review the economic literature on the relationship between ageing, morbidity and care expenditure, followed by a discussion of Swedish LTC provision. In the third section, we present the dataset and discuss our empirical strategy. Our results are then presented along with an estimate of the budget effects if life expectancy were to hypothetically increase by one year. In addition, we provide long-term projections based on demographic projections from the Swedish statistical office. The final section summarises our findings and the subsequent policy implications.

2.2 Healthcare Expenditures and Ageing Populations

Over the last decade, much attention has been devoted to the so-called 'red herring' hypothesis, according to which care costs are unrelated to age once remaining lifetime or proximity to death is controlled for. The literature on this topic can be divided into studies investigating curative care and those investigating LTC provision. Most studies use individual-level data, although there are contributions using aggregated data.

In their widely acknowledged study, Zweifel et al. (1999) investigate the red herring hypothesis using Swiss data. They show that the impact of age on healthcare costs decreases once TTD is taken into account. During the last two years of life, an individual's actual age seems to be completely irrelevant. This leads to their conclusion that age is not necessarily an important determinant of healthcare expenditures. Subsequent studies address methodological issues and provide further empirical evidence. In an excellent literature review, Payne et al. (2007) provide a

picture of healthcare expenditures and their interplay with ageing, morbidity and death. They conclude that, although the impact on cost predictions may be small, using both age and TTD in expenditure models provides a clear advantage over simple age-based models by helping to evaluate which services can be provided most efficiently. A prominent contribution by Felder et al. (2010) suggests that future increases in healthcare expenditures are more likely to be caused by changes in medical technology than further ageing of the population. Wong et al. (2011) analyse disease-specific hospital expenditures and find that, while age is a relevant determinant of healthcare expenditures, its effect is modest compared to that of TTD. In general, recent evidence based on individual-level data suggests that end of life morbidity (captured by TTD) is the main predictor of healthcare expenditures, while ageing itself is of minor relevance.

There are also studies into the red herring hypothesis as it relates to LTC expenditures. Werblow et al. (2007) used panel data from Switzerland on healthcare and LTC provision and found that most components of care expenditure are driven by TTD rather than age. Forma et al. (2007) find that, in Finland, hospital stays increase in the last months of life and demand for public LTC increases strongly in the 2 years prior to death. In a study using a Swedish sample, Larsson et al. (2008) find that home help provision is influenced by age and not TTD, while institutional and hospital care are much more influenced by end of life morbidity than age. Their results also suggest an increased transition from home- and community-based care to institutional care during the last year of life. Using a Dutch dataset on the use of institutional LTC and home care, de Meijer et al. (2011) differentiate between causes of death and analyse the impact of morbidity. Once morbidity and disability are controlled for, age remains relevant, while TTD becomes insignificant. They conclude that TTD cannot causally affect care expenditures and might itself be interpreted as a red herring: as a simple proxy for morbidity and disability. Hence, evidence on LTC provision based on individual-level data also suggests that end-of-life morbidity is an important driver of LTC provision, but findings for ageing are mixed. There also seem to be marked differences between the provision of institutional and home LTC, as institutional care is in most cases only provided at the very end of life.

Although individual-level data is usually preferred in such analyses, there are also some issues that are better addressed at an aggregated level. Besides the fact that aggregated data is often more readily available, evidence based on it is usually more representative, particularly if it is based on census data containing information about all individuals living within a certain unit of observation.

Few studies into the red herring hypothesis use aggregated data. Palangkaraya and Yong (2009) try to tackle this issue using panel data from 22 OECD countries. They find the proportion of people aged 65 and older (reflecting age) does not explain healthcare expenditures once mortality and other factors are controlled for. In a study on the EU-15 countries, Bech et al. (2011) focus on the relationship between population demographics, mortality, life expectancy and health expenditures, finding only short-term effects of ageing on healthcare expenditures. Their findings also suggest that past rather than present mortality rates determine healthcare expenditures, which they explain as political lag. This is a conceptually important distinction from the interpretation of mortality in studies using individual-level data, where TTD (capturing individual mortality) is assumed to be directly associated with higher healthcare expenditures. Breyer et al. (2015) analyse panel data including German sickness fund members at the cohort level. They distinguish between the effects of ageing, contemporary mortality and 5-year survival rates. They interpret the survival rates as capturing changes in longevity. As in Bech et al. (2011), these survival rates are prospective measures, whereas yearly mortality rates are retrospective. The authors find both mortality rates and ageing are positively correlated with healthcare costs. Karlsson and Klohn (2011) use simple mortality rates to capture the impact of TTD on expenditures for overall social care for the entire Swedish population. Their findings suggest that age is a much more relevant determinant of Swedish social care expenditures than mortality. In a recent study, van Baal and Wong (2012) evaluate the extent to which including TTD influences forecasts of macro-level healthcare expenditures. They compare different scenarios based on their estimates and find that including mortality does not decrease predicted expenditures.

Results for the impact of ageing and TTD on care expenditures are strongly heterogeneous in studies using aggregated data. This might be explained by not just the data but also diverging empirical approaches. Furthermore, unlike studies analysing healthcare expenditures, studies focusing on LTC at the individual level show that TTD is a more robust determinant of care costs.

There is a dearth of research on the red herring analysis using aggregated data. Freely available datasets of official statistics are attractive and potentially helpful to policy-makers. However, in order to provide more conclusive evidence on the red herring hypothesis using aggregated data, several gaps need to be addressed. For example, the use of raw mortality rates is conceptually unconvincing, since they do not represent the aggregated equivalent of individual proximity to death. By focusing on income elasticities of healthcare expenditures, Getzen (2000b) highlights

another difference between studies implemented on the micro and macro levels. He argues that the unit of an analysis is very important, because healthcare tends to be a necessity at the individual level and a luxury at the macro level. He says a risk pooling group largely eliminates issues such as ability to pay, while differences between groups are largely determined by differences in funding structure. As the aim of risk pooling is to diversify risk within a group, the link between risk and expenditures may no longer be observable at the aggregated level. In addition, most studies focus exclusively on aggregate healthcare expenditures, making LTC provision increasingly relevant.

We fill this gap by providing the first study to evaluate the red herring hypothesis in the market for LTC provision at an aggregate level. First, we combine the contemporary and future mortality rates of our population of interest to use the probability of dying as a retrospective measure to control for TTD at an aggregate level. Accounting for future mortality is crucial when decomposing LTC expenditures into age and TTD effects at the aggregated level, since death-related morbidity is not exclusively restricted to the last year of life. Since the overall aim of redistributing resources between Swedish municipalities is to compensate for differences in tax bases and cost structures without altering service quality (Karlsson et al., 2010), we consider the problems raised by Getzen (2000b) when working with aggregated macro data to only be a minor issue for our analysis. Differences in patient characteristics that are beyond the control of local government are not expected to be associated with a region's ability to pay if variations in costs are eliminated by national government. Hence, Sweden provides a favourable institutional background for our empirical analysis (Getzen, 2006).

2.3 LTC Provision in Sweden

The main goal of the Swedish LTC system is to provide high quality services to every resident according to their needs. While county councils are responsible for healthcare provision (e.g., hospitals and health centres), municipalities are responsible for all other aspects of care, including social care, institutional care and home nursing. Although services can be supplied directly by a municipality or by a private health and social care provider, local authorities remain responsible for funding them (Fukushima et al., 2010). Directly elected politicians decide on the supply of LTC, as well as raising the revenues necessary to cover expenditures. The main source of funding is local income taxes, and out-of-pocket payments are of minor importance

(4 % of total costs,(Colombo et al., 2011)). The national government lays down general principles and responsibilities for social care in law and monitors care home quality. It also redistributes funds to create equal conditions for the provision of LTC in all parts of the country, despite immense differences in both need and local tax bases. This equalisation is based on local income, expected costs and other structural disadvantages. In 2008, the national government transferred 58 billion SEK (€5.9 bn.) to local authorities, 17.6 % of total revenue (Karlsson et al., 2010). In order to receive LTC services, an application must be made to the local authority. An evaluator then interviews the potential recipient and family members to determine the extent of required support, and whether the services can be provided as domiciliary care. This evaluation is based on restrictions to daily living activities (ADL). Institutional care is viewed as a last resort (Fukushima et al., 2010).

Although the market share of private providers has increased over the last two decades in Sweden (e.g., from 5.4 to 13.7 % of nursing home slots (Socialstyrelsen, 2008)), nearly all formal LTC services are still funded and monitored by local authorities.

2.4 Data and Empirical Strategy

To evaluate the red herring hypothesis in the context of expenditures on LTC services, we use administrative data collected by the Swedish National Board for Health and Welfare. This data is provided at the municipality level, which allows us to exploit the panel structure of our data. Since we can analyse high-quality data aggregated at the regional level[2] within an institutional framework that redistributes wealth to meet local needs, there are very good opportunities to evaluate the relationship between ageing and LTC without some of the typical confounders.

In some counties, municipalities have also taken over responsibility for healthcare provision; unfortunately, data from these counties is ambiguous in that it does not specifically identify LTC and healthcare expenditures. As the focus of our analysis is on LTC expenditures, we exclude such observations. The remaining municipalities cover more than half of the Swedish population, including the counties of Blekinge län, Gävleborgs län, Kalmar län, Norrbottens län, Stockholms län, Värmlands län, Västerbottens län, Västernorrlands län and Östergötlands län. In addition, in parts of Dalarna län, Jönköpings län and Västmanlands län, around 10 % of the municipal-

[2]Data on LTC expenditures is available at http://www.scb.se.

ities provide healthcare services. Although comparing variables for the two groups[3] shows they are very similar, we do not use such observations in our analysis. As the quantity of healthcare services offered over time cannot be distinguished from LTC provision for these regions, this would clearly be a potential confounder in our analysis, because our fixed effects estimator uses deviations from the municipalities' means over time for identification. Hence, we exclusively use the data on LTC provision, representing more than half of the population (capturing 4.9 out of 9.2 million individuals).

Our main units of observation are 152 out of a total of 290 Swedish municipalities[4]. Our analysis focuses on the sub-population of inhabitants aged 65+, since only this group is eligible for LTC. Therefore, our main dependent variable captures publicly financed, overall LTC expenditures[5] divided by the average population aged 65 or older, but we also provide separate regression estimates for institutional and domiciliary LTC costs. Since all variables are expressed in terms of averages per 65+ inhabitant, all estimated parameters correspond to their individual-level equivalents. As can be seen in Table 2.1[6], the average costs for overall LTC expenditures are around SEK 57,400 (€6,400) per capita. Institutional and domiciliary expenditures do not completely account for total LTC, as the figure includes additional services which cannot directly be attributed to one of the two categories, such as preventive care.

To account for TTD in our model, we consider contemporary and future mortality rates among the 65+ population. Thus, we assume a high mortality rate among the 65+ population is related to a high level of care and is therefore also positively correlated with LTC expenditures. According to the red herring hypothesis, the mortality rate for the following period should also be correlated with higher LTC use, since individuals dying in the next year also have a higher probability of using LTC services today. Hence, we define the TTD variables as

$$TTD_{it}^{a} = (1 - (1 - mrt_{it}^{a}) \times (1 - mrt_{i,t+1}^{a+1})) \times \frac{N_{it}^{a}}{N_{it}^{65+}} \tag{2.1}$$

which can be interpreted as the probability of dying within two years for people

[3]See Appendix 2.A1

[4]Data provided by Statistics Sweden can be downloaded at http://www.ssd.scb.se.

[5]As already mentioned, some 4 % of total costs are covered by user charges. These are included in our cost variables to reflect the actual total costs of LTC for each individual in the publicly funded system.

[6]For a detailed variable description, see Appendix 2.A2

Table 2.1 Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Total	57.4	9.418	30.995	94.41	1,589
Inst	37.532	8.609	14.771	70.396	1,589
Dom	17.536	5.659	4.258	39.3	1,589
age6569	0.267	0.034	0.203	0.457	1,589
age7074	0.236	0.016	0.19	0.281	1,589
age7579	0.21	0.017	0.142	0.264	1,589
age8084	0.158	0.019	0.084	0.21	1,589
age8589	0.089	0.014	0.045	0.132	1,589
age9094	0.033	0.007	0.012	0.06	1,589
age95100	0.007	0.002	0	0.017	1,589
mrt	0.052	0.007	0.031	0.089	1,589
mrtL1	0.054	0.007	0.03	0.082	1,589
TTD65	0.102	0.012	0.065	0.147	1,589
TTD6569	0.008	0.002	0.001	0.017	1,589
TTD7074	0.011	0.003	0.003	0.023	1,589
TTD7579	0.017	0.004	0.005	0.034	1,589
TTD8084	0.024	0.004	0.007	0.041	1,589
TTD8589	0.023	0.005	0.006	0.041	1,589
TTD9094	0.014	0.003	0.005	0.029	1,589
TTD95100	0.005	0.002	0	0.011	1,589
medinc65_08	156.551	20.648	100.984	240.874	1,589
privcare	0.061	0.133	0	0.946	1,435
wom65	0.557	0.018	0.498	0.633	1,589
density	157.767	537.992	0.2	4307.8	1,589
rightwing	0.337	0.106	0.105	0.764	1,589
taxrate	0.214	0.013	0.149	0.231	1,589
mrtl1	0.052	0.007	0.032	0.089	1,589
lifexp	18.424	1.035	14.015	22.52	1,589

Summary statistics for the dependent and explanatory variables of our baseline specification (not sex specific).

in a specific age group (cf. Wilmoth et al. (2007)). More specifically, our overall TTD variable, TTD_{it}^{65+}, is defined as the probability of dying within two years for an individual aged 65+, whereas the age-specific TTD is the probability of dying within two years multiplied by the proportion of the 65+ population within that specific age range, *a*. Hence, it gives the weighted probability of dying within two years for a given age group. The mortality rates *mrt* are calculated as the number of older people in a specific age group, *a*, who died within a year, divided by the number of people, N_{it}^{a}, in the age group alive at the beginning of the same year. Using the population at the beginning of each period is important, since the number of deaths in a given period negatively affects the number of inhabitants within a municipality. The next year's mortality rate is calculated with respect to the initial age group, *a*, which we now (with some abuse of notation) call $a + 1$, as we need to account for the fact that individuals from age group *a* are one year older in period *t+1*. The final term in equation (2.1) is used for rescaling the age- and sex-specific TTD variables to the 65+ level, as the pure survival rates resulting from the first two terms were calculated for the group-specific population numbers, not reflecting their actual relevance, given the number of older (65+) people in each municipality. In this way, we can provide a reliable retrospective measure for TTD to control for an individual's proximity to death (here: probability of dying within two years) at an aggregated level. Exploring age-specific mortality in future periods is an important distinction from existing macro studies using aggregated data (e.g., Bech et al. (2011), Breyer et al. (2015)), and important if end-of-life morbidity is not restricted to the last year of life. In our baseline model, *a* simply represents the total 65+ population, but group-specific TTD variables created with respect to age group *a* are also used later in our analysis (e.g., TTD_{it}^{6569} and sex-specific TTD variables [7]).

It could be argued that accounting for the probability of dying within the next two years is insufficient, since dying in more distant periods might also generate extra costs of dying. Thus, we also derived the TTD variable based on three periods. The results suggest that using three periods does not increase LTC costs significantly. As using more distant information decreases the number of observations available for analysis, what follows relies on the TTD variable based on two periods. Another important factor in the choice of number of periods is immigration: As we are not observing individuals but municipalities, there is the potential problem that elderly people may move from one municipality to another. Using more distant periods to

[7] The pattern of lower probabilities of dying for higher age groups, as seen in Table 2.1, is simply a result of the variables being rescaled as described above.

calculate the TTD variable would increase measurement error if some individuals changed municipalities at the end of their life. However, we regard this issue as being of minor empirical relevance, as we know that older individuals are less likely to change their municipal residency[8].

The other main explanatory variables are those capturing the age structure amongst the old in each municipality. These variables are defined as the number of older people, measured in intervals of five years, divided by the number of people aged 65+. Again, these variables are measured at the beginning of each year. As expected, the share of individuals decreases with the specific age group, from 27 % to less than 1 % on average.

Other variables used for robustness checks include the median incomes (1,000 SEK) of people aged 65+, the share of private LTC provision, life expectancy for people aged 65+, the lagged mortality rate, local tax rates, the centre-right parties' share of all seats in the town council, population density and the share of women in the 65+ population. To account for inflation, we standardised all monetary variables for 1998 to 2008 according to the Swedish price index, expressing them in 2008 SEK. Due to missing values in the data, our final sample is restricted to 1,589 observations.

We assume a flexible relationship between our explanatory variables (age groups and TTD_{it}^{a}) and the dependent variable (care costs per 65+ population), as we allow ageing to have a heterogeneous impact on LTC costs. Endogeneity may be an issue in this analysis, and the possibilities to adjust for it are limited. While the age distribution at the beginning of the year is clearly predetermined, it may be correlated with unobservable characteristics that explain LTC expenditures. The TTD variable is also problematic in this regard. Hence, we rely on a fixed effects estimator[9], assuming that possible confounders in our baseline specification are constant over time. The functional relation is

$$y_{it} = X\beta + \mu_i + \nu_{it} \tag{2.2}$$

where β is a vector of length $k \times 1$, and X is an $N \times K$ matrix containing a constant, the age variables and other controls depending on the specification. μ_i is a municipality-specific error term. In addition to the baseline specification, we provide estimates with other covariates included to evaluate the robustness of our findings. We also estimate the impact of age and TTD on LTC costs, allowing for age-specific

[8]The 65+ population changes its municipality with a probability of around 0.011. This drops to 0.007 for the 80+ population (2006; Centralbyrån (2009)).

[9]A Hausman test supports the hypothesis that the consistency of a random effects estimator can be rejected.

TTD effects. This specification is saturated since, apart from municipality and time effects, it controls for all possible combinations of our age categories and TTD. Hence, the functional form of the relationship between independent and dependent variables is less of an issue. In another specification we calculate age and TTD with respect to the municipality-specific distribution of sex, as done in earlier studies (e.g., Felder et al. (2010), Breyer et al. (2015)). Since we observe expenditures at the municipality rather than individual level, we estimate specifications using the same dependent variables as before, and 'interact' the explanatory variables capturing age and TTD with indicators for the region-specific distribution of men and women. To allow for a straightforward interpretation of this specification and avoid assumptions about sex-specific costs in the control group, we exclude the constant from this specification. For this reason, coefficients for the share of 65- to 69-year-old males and females are also provided here. To account for heteroskedasticity, all estimates use weighted least squares. The weights reflect the inverse relation between the variance of the outcome variable and the size of the population.

2.5 Findings

2.5.1 Baseline Specification

Table 2.2 shows the results of our fixed effects specifications. The table provides a comparison between the impact of ageing on total, institutional and domiciliary LTC expenditures when TTD is omitted and included.

For overall LTC costs, the coefficients for the older age groups are positive and significant. This suggests that age-related costs are not a major issue in these younger cohorts.

When TTD is excluded (column 1), the coefficient for 85- to 89-year-old people indicates that an individual in this age group incurs an increase in total LTC expenditures of SEK 103,000 (€10,000) per year (in addition to the average for an individual aged 65 to 69, captured by the constant). In accordance with the red herring hypothesis, the coefficients for the older age groups decrease once the probability of dying within two years is included in the model.[10] However, the effect of including TTD is modest. Thus, our results support the existence of a red herring, but age

[10]We also estimate specifications which include both contemporary and next year's mortality. Wald tests do not reject the null hypothesis of the equality of both coefficients, suggesting they control for the same mechanism influencing LTC costs. Therefore, we regard aggregating both variables as an appropriate way to control for the overall effect – TTD^{65+}.

itself seems to have a strong impact on LTC expenditures, even after controlling for mortality.

Table 2.2 Impact of TTD

	(1) Total	(2) Total	(3) Inst	(4) Inst	(5) Dom	(6) Dom
age7074	-10.61	-13.70	13.83	9.62	-13.53	-12.66
	(22.83)	(22.43)	(26.20)	(25.92)	(21.59)	(21.76)
age7579	-14.22	-20.16	-25.45	-33.53	6.78	8.44
	(22.81)	(22.62)	(31.45)	(31.59)	(23.88)	(24.21)
age8084	27.52	19.26	23.77	12.53	10.41	12.73
	(26.69)	(27.00)	(29.91)	(29.83)	(21.55)	(22.12)
age8589	102.81***	87.06***	37.57	16.13	54.20**	58.62**
	(29.33)	(29.55)	(27.18)	(28.15)	(22.08)	(23.21)
age9094	214.69***	188.17***	105.37*	69.29	98.68**	106.11**
	(40.64)	(39.50)	(54.30)	(55.63)	(42.42)	(44.54)
age95100	273.74***	248.44***	93.33	58.89	187.11*	194.20*
	(90.29)	(90.19)	(125.60)	(125.63)	(102.70)	(104.10)
TTD65		60.63***		82.50***		-17.00
		(19.21)		(23.08)		(19.12)
Constant	37.78***	37.73***	23.86	23.79	9.80	9.82
	(13.50)	(13.29)	(16.52)	(16.24)	(12.38)	(12.37)
Observations	1,589	1,589	1,589	1,589	1,589	1,589
R^2	0.557	0.561	0.234	0.244	0.452	0.453

Fixed effects estimates for our three cost categories, including and excluding TTD. Columns 1 and 2 show the coefficients for total LTC costs, 3 and 4 for institutional LTC and 5 and 6 for domiciliary LTC. Year dummies are included, and the regressions are weighted by the square root of a municipality's average 65+ population. The unit of observation is a Swedish municipality for the period 1998-2008. Standard errors are clustered at the municipality level. * p<0.10, ** p<0.05, *** p<0.01

Separate estimates for institutional and domiciliary LTC costs reveal differing patterns for most age groups, suggesting ageing is the most relevant predictor for domiciliary care, whereas TTD is much more relevant for institutional LTC costs. For institutional care, we also find the effects of ageing are offset by the inclusion of TTD. The differing signs of the coefficients for TTD might be interpreted as an indicator of age-related switching behaviour from domiciliary into institutional care. This finding is in line with Larsson et al. (2008), who argue there is a transition process from home to institutional LTC at the end of life.

Overall, the estimated morbidity effects in all scenarios but the institutional care are dwarfed by the increase in expected costs at higher ages. Thus, our conclusion

here differs from that of Larsson et al. (2008), who find that TTD is much more important than age in Sweden. However, their study is implemented on the individual level and does not use monetary equivalents of LTC utilisation.

To check whether our estimates suffer from omitted variable bias, we further consider a specification that takes into account other variables that might be potential determinants of LTC expenditures. The sociodemographic variables are the median income of the 65+ population, population density and the share of females in the 65+ population. To account for potential variation in the supply side of LTC services, we add the share of private LTC providers. In addition, municipalities' health and LTC policy might vary over time and thus not be accurately captured by municipality fixed effects. We add the local tax rate and the centre-right parties' share of all local government seats to account for this. As studies emphasise that changes in life expectancy are an important determinant of healthcare provision (e.g., Bech et al. (2011), Breyer et al. (2015)), we add the local life expectancy of 65-year-old people to our model. We include the lagged mortality rate, which Bech et al. (2011) find is relevant. Including these variables allows us to check whether the estimated coefficients are robust or simply reflect a partial correlation with other important variables.

We find the age coefficients (shown in Table 2.3) are slightly higher than for our baseline model. The median income of the elderly population, local tax rate and lagged mortality rate are significantly correlated with LTC expenditures. The positive correlation with income is remarkable, since equalisation grants from the national government are provided to compensate for differences in local needs. Thus the coefficient might still capture an income effect on demand. The tax rate and lagged mortality rate capture municipality-specific LTC provision. The relevance of lagged mortality is in line with empirical evidence on the healthcare market (Bech et al., 2011). Bech et al. (2011) argue that it takes time for shifts in demand for healthcare services to be incorporated into public funding. Interestingly, increasing life expectancy did not affect our estimates, once conditioned on TTD. However, as we are exclusively interested in the effect of ageing on LTC expenditures, the inclusion of potential confounders can be problematic, because we do not know the underlying mechanism that relates these variables to LTC provision. Hence, later in this study we use the estimates of our baseline specification to calculate economic implications, keeping in mind that these projections are conservative.

Table 2.3 Robustness estimates

	Full sample	Full sample	Restricted	Restricted
	Total	Total	Total	Total
age7074	-13.70	-6.39	-3.87	-0.75
	(22.43)	(21.20)	(21.52)	(20.19)
age7579	-20.16	-12.13	-13.34	-9.18
	(22.62)	(23.70)	(21.96)	(22.93)
age8084	19.26	41.38	27.05	43.65
	(27.00)	(26.17)	(29.56)	(28.05)
age8589	87.06***	120.30***	84.96***	118.14***
	(29.55)	(27.29)	(31.15)	(28.96)
age9094	188.17***	216.78***	185.12***	213.81***
	(39.50)	(39.61)	(39.22)	(40.92)
age95100	248.44***	320.88***	243.50***	310.59***
	(90.19)	(92.16)	(90.84)	(89.95)
TTD65	60.63***	42.35**	49.64**	34.59
	(19.21)	(21.15)	(19.64)	(21.69)
medinc65_08		0.23***		0.23***
		(0.07)		(0.07)
wom65		-6.81		-0.10
		(28.99)		(29.15)
density		-0.01		-0.01
		(0.01)		(0.01)
rightwing		-0.95		-3.88
		(6.59)		(6.91)
taxrate		174.49***		165.02***
		(51.48)		(52.40)
mrtl1		55.80***		56.15***
		(20.83)		(20.88)
lifexp		-0.23		-0.20
		(0.14)		(0.14)
privcare				-0.93
				(2.65)
Constant	37.73***	-40.27	34.37**	-42.60
	(13.29)	(27.00)	(13.53)	(26.57)
Observations	1,589	1,589	1,435	1,435
R^2	0.561	0.585	0.496	0.524

Fixed effects estimates with total LTC expenditures as the dependent variable. We estimate our baseline specifications with and without other control variables that might determine LTC expenditures. As we do not have information on private LTC provision for all our municipalities, we performed the robustness check twice to determine whether excluding observations with missing values might change our estimates. Year dummies are included and the regressions are weighted by the square root of each municipality's average 65+ population. The unit of observation is a Swedish municipality for the period 1998-2008. Standard errors are clustered at the municipality level. * p<0.10, ** p<0.05, *** p<0.01

2.5.2 Age-Specific TTD Effects

In addition to the specifications above, we now analyse the red herring hypothesis in more detail by testing whether TTD impacts LTC expenditures differently for different age groups. This is the analogue to a fully interacted specification in an analysis using micro data, i.e., this specification reflects the spirit of a saturated model, as we allow for all combinations of our age variables and TTD. If TTD effects vary with age, the age coefficients might falsely pick up TTD-related variations in LTC costs. We therefore estimate our baseline specification for overall, domiciliary and institutional costs with age-specific TTD variables, i.e., instead of the single TTD measure TTD^{65+}, we use group-specific TTD^{a} variables.

The results in Table 2.4 suggest that LTC costs differ both between age groups as well as between domiciliary and institutional care. Again, compared to our baseline model, the inclusion of age-specific TTD variables increases the impact of age for the older age groups. The comparison between domiciliary and institutional care supports the hypothesis, mentioned in the last section, that the negative coefficient of TTD for domiciliary LTC is driven by switching between care settings at the end of life. This substitution between domiciliary and institutional LTC is mainly relevant for people aged 90 to 94. However, the negative coefficient for domiciliary care is much higher than the positive corresponding coefficient for institutional LTC, i.e., for the oldest individuals, the shift to institutional care yields a decrease in total cost. When we look at the TTD coefficients in the specification including institutional care, we find the strong positive effect of our baseline estimate to be mainly driven by the relatively young, the 70- to 74-year-olds, for whom age is not a relevant determinant of LTC expenditures. The estimates strongly suggest an age-specific impact of TTD between institutional and domiciliary LTC expenditures. But once again, the age coefficients seem to be most relevant for domiciliary care.

In summary, ageing is again the main driving force behind LTC expenditures, even if we allow for age-specific TTD influences.

2.5.3 Sex-Specific TTD Effects

Although sex is already taken into account via our robustness checks, we also identify the parameters of interest separately for women and men. Hence, relative to the estimated specifications above, we further relax the assumption of homogeneous relationships between our age variables and LTC provision.

We again provide different specifications for our LTC categories. First, we run a

Table 2.4 Age-specific TTD effect

	(1) Total	(2) Inst	(3) Dom
age7074	-21.66	6.20	-17.38
	(22.59)	(26.26)	(22.30)
age7579	-21.26	-29.80	2.56
	(23.68)	(32.01)	(24.11)
age8084	18.43	15.44	7.21
	(27.99)	(30.92)	(22.45)
age8589	103.36***	23.44	64.46***
	(31.25)	(29.96)	(23.72)
age9094	232.77***	84.17	146.91***
	(45.72)	(65.93)	(50.53)
age95100	334.09***	84.20	272.96**
	(119.98)	(165.16)	(133.08)
TTD6569	64.54	114.29	-50.62
	(68.92)	(69.33)	(60.61)
TTD7074	210.65***	179.63***	29.40
	(59.64)	(60.08)	(51.67)
TTD7579	86.23	55.92	40.48
	(52.49)	(48.92)	(35.89)
TTD8084	88.32**	87.24*	13.86
	(39.99)	(49.85)	(40.38)
TTD8589	-9.15	59.43	-51.64
	(42.31)	(53.12)	(36.10)
TTD9094	-45.28	48.53	-120.58**
	(57.09)	(64.69)	(55.63)
TTD95100	-57.13	47.31	-120.84
	(121.94)	(133.91)	(109.58)
Constant	37.38***	22.30	11.38
	(13.26)	(16.19)	(12.27)
Observations	1,589	1,589	1,589
R^2	0.566	0.246	0.457

Fixed effects estimates for our three cost categories when age-specific TTD is accounted for. Column 1 shows the coefficients for total LTC costs, 2 for institutional and 3 for domiciliary LTC. Year dummies are included and the regressions are weighted by the square root of each municipality's average 65+ population. The unit of observation is a Swedish municipality for the period 1998-2008. Standard errors are clustered at the municipality level. * p<0.10, ** p<0.05, *** p<0.01

fixed effects regression using only sex-specific age variables[11], we then include sex-specific TTD variables, followed by life expectancy (which was also found to be different for women and men in the descriptive statistics). The final regression, shown in columns 4-5 of Table 2.5, includes TTD variables that allow for heterogeneity in both sex and age. Unlike the specifications in the sections above, we excluded the constant from these estimates, introducing the share of 65- to 69-year-old men/women as additional covariates. This is because the dependent variable cannot be defined separately for men and women. Hence, in order to avoid the same constant for men and women, we account for age effects of both sexes in these estimates.

For total LTC expenditures we find a strong sex-specific pattern to the age variables, as seen in Table 2.5. Although age is again positively related to LTC expenditures for all individuals aged 80+, this effect is overall much stronger for women than for men. The overall TTD effects in our next specification (column 2), again offsetting the effect of ageing, shows that the average TTD effect is mostly driven by men. Introducing sex-specific life expectancy (column 3) does not change these findings much. Looking at joint age- and sex-specific TTD effects, we again find the highest TTD effects for people aged 70-74. Here offsetting age by including TTD is found to be of minor relevance.

When focusing on institutional and domiciliary LTC separately (Appendix 2.A4 and 2.A5), we find the TTD effect for individuals aged 70-74 to be driven by institutional care. Once again, the positive TTD effects are not prevalent when focusing on domiciliary care; here TTD for women is strongly negatively associated with LTC expenditures. Both effects seem to offset themselves in the aggregated variable total LTC. The fact that TTD is only negatively related to domiciliary LTC for older women is an interesting finding. This indicates that switching behaviour between domiciliary and institutional care at the end of life mostly relates to women. We suppose that this can be attributed to the relevance of informal care as well. Sex-specific differences in informal care are discussed by Paraponaris et al. (2012), who find that being female is a strong determinant of receiving formal care. Our age-specific decomposition shows that the effect is highest for the 65- to 69-year-old and 85- to 94-year-old women. The same age groups show a positive TTD effect when using institutional care as the dependent variable. We find this pattern for the age-specific TTD variable and the 65+ TTD variable. Again, the small change in age effects due to the inclusion of TTD casts doubt on the relevance of the red herring hypothesis, i.e., excluding the probability of dying does not change the economic relevance of

[11]Descriptive statistics are provided in Appendix 2.A3

Table 2.5 Sex-specific total LTC expenditures

	(1) Total	(2) Total	(3) Total	(4) Total	(5) Total
				Men	Women
age6569m	42.61*	39.94*	42.38*	42.89*	
	(23.89)	(23.90)	(24.34)	(24.40)	
age7074m	35.79	25.67	27.93	12.80	
	(27.97)	(28.38)	(28.43)	(29.71)	
age7579m	-3.27	-17.48	-14.30	-25.97	
	(32.28)	(33.75)	(33.83)	(34.84)	
age8084m	44.65	23.11	26.79	29.30	
	(43.34)	(47.14)	(46.98)	(48.27)	
age8589m	87.70*	54.24	59.26	96.49	
	(52.92)	(55.09)	(55.28)	(62.26)	
age9094m	245.58***	193.70**	199.69**	254.97**	
	(86.05)	(89.14)	(89.86)	(108.85)	
age95100m	121.19	106.19	115.76	144.31	
	(186.79)	(185.84)	(188.37)	(281.03)	
age6569w	27.85	32.10	34.79		33.01
	(25.45)	(25.67)	(26.04)		(27.05)
age7074w	22.48	25.38	28.37		21.74
	(29.46)	(29.84)	(30.41)		(30.51)
age7579w	45.53*	45.49*	48.92*		49.86*
	(24.21)	(24.68)	(24.92)		(25.58)
age8084w	73.56**	74.04**	78.29**		73.02**
	(30.09)	(31.35)	(31.15)		(33.85)
age8589w	164.35***	158.85***	163.87***		159.19***
	(36.02)	(37.61)	(37.70)		(41.38)
age9094w	249.44***	233.13***	238.96***		270.40***
	(50.90)	(50.24)	(50.86)		(61.25)
age95100w	393.19***	372.40***	381.66***		460.86***
	(103.93)	(108.18)	(108.93)		(146.18)
TTD65m		79.59**	73.42**		
		(34.23)	(35.86)		
TTD65w		45.77	32.01		
		(31.50)	(34.67)		
lifexpm			-0.03	-0.00	
			(0.09)	(0.10)	
lifexpw			-0.09		-0.05
			(0.10)		(0.10)
TTD6569s				45.37	56.58
				(90.92)	(121.90)
TTD7074s				216.87***	209.90**
				(79.23)	(92.21)
TTD7579s				140.42**	27.74
				(66.22)	(73.38)
TTD8084s				83.30	68.55
				(65.17)	(59.55)
TTD8589s				-41.86	26.22
				(66.21)	(60.56)
TTD9094s				-22.30	-72.66
				(118.99)	(72.70)
TTD95100s				75.38	-95.09
				(271.57)	(156.61)
Observations	1,589	1,589	1,589	1,589	

Fixed effects estimates for total LTC expenditures when age-specific control variables are included. We included sex-specific coefficients for the 65-69 age group as we do not rely on a constant in this specification. Column 1 accounts for age, column 2 includes TTD and column 3 uses life expectancy. In columns 4 and 5 (coefficients taken from a single regression), we allow for age- and sex-specific TTD effects on total LTC expenditures. Year dummies are included, and the regressions are weighted by the square root of each municipality's average 65+ population. The unit of observation is a Swedish municipality for the period 1998-2008. Standard errors are clustered at the municipality level. * p<0.10, ** p<0.05, *** p<0.01

ageing very much.[12]

Overall, we find the age coefficients, especially for women, are highly significant and economically relevant when focusing on total expenditures, whereas some become statistically insignificant for domiciliary and institutional costs. As we again find a sex-specific shift between the two types of services, but lack reliable information about future changes to the distribution of men and women, we do not use these estimates to project care expenditures in the following section.

2.5.4 Implications

To illustrate how our findings can be used to project costs, we first calculate the expenditures arising due to a hypothetical increase in the life expectancy of the 65+ population of exactly one year. We show how costs change for different estimated specifications. Second, we use population projections from the Swedish Statistical Office for the next 100 years to further project the financial consequences of expected changes in age distribution on total LTC expenditures, based on our estimates. These projections are inversely weighted by the total population (and the 15- to 64-year-old workforce), to reveal the importance of differences in the potential to contribute to funding LTC services for the elderly.

The following equation shows the implied changes in expenditures driven by increased life expectancy[13]:

$$\triangle LTC_i = \begin{cases} \sum_{k=65}^{\infty} \left(\alpha_j + \beta_j^k\right)\left(s_{2008}^{k-1} - s_{2008}^{k}\right) \; with \; j = i & \text{,for i=1,2} \quad (2.3a) \\ \sum_{k=65}^{\infty} \left(\left(\alpha_j + \beta_j^k\right)\left(s_{2008}^{k-1} - s_{2008}^{k}\right) + \delta_j^k \left(TTD_{2008}^{k-1} - TTD_{2008}^{k}\right)\right) \; with \; j = i & \text{,for i=3} \quad (2.3b) \\ \alpha_{j=1} & \text{,for i=4} \quad (2.3c) \end{cases}$$

where we calculate the changes for total, institutional and domiciliary LTC costs ($\triangle LTC$) and distinguish between four different scenarios, i, by combining our estimates with the Swedish age structure and age-specific mortality from 2008. Since this information is provided for every age cohort, the overall cost projection can be

[12]To evaluate the extent to which the coefficients of sex-specific TTD are reliable and not a result of colinearity due to including so many TTD variables, we ran single regressions including only a single TTD variable. One explanation for the problem could be comorbidity of very old people, which would not allow TTD effects to be identified separately for both sexes. However, these estimates prove the findings for the entire specification are very similar, both in terms of economic relevance and statistical significance.

[13]The formula for cost increase per life year is provided in Appendix A2.6.

written as the sum of age-specific projections from ages 65 to infinity, but ends at k=105 in our setup.

We differentiate between a naive demographic extrapolation (i.e., not taking TTD into account), two scenarios where TTD (overall and age-specific) is accounted for, and a pure red herring scenario (i.e., all age-related costs are costs of dying). j indicates which parameter estimates are used for each scenario, i. If j=1, we use coefficients from a regression where only age is controlled for (odd columns in Table 2). If j=2, then TTD^{65+} is included, whereas j=3 indicates the inclusion of the seven age-specific TTD^{a} variables into the regression. If j=4, we again use the parameter estimate of the specification controlling only for age. In this pure red herring scenario, we assume that all relevant costs are captured by a constant, although we accounted for age in the specification. α represents our estimated constants, β the age coefficients (and is therefore fixed within five-year intervals). s^{k}_{2008} are year-specific survival rates based on Swedish life table estimates, which we use to model a hypothetical increase in life expectancy of exactly one year. We assume that each individual is one year 'younger' (in terms of mortality) in the counterfactual situation, i.e., the survival rate of age cohort k in period t equals s^{k-1} in year $t+1$. TTD^{k}_{2008} are TTD variables calculated for individual ages $k = \{65,..,100\}$.

The TTD variable is only included in the formula for the specification controlling for age-specific TTD effects, as they only occur once in an individual's life and are therefore only influenced by a change in life expectancy if they vary with age. To quantify the implications of our estimates, we exclusively rely on the size of our estimated parameters and do not use statistical significance.

Figure 2.1 shows the implied change in lifetime LTC costs (from the current level of SEK 947,000 in scenario 1 for total costs), and Figure 2.2 shows the corresponding increase in average costs per person per life year (assuming a stationary population). Although our results imply lower cost increases for total and institutional LTC costs once TTD^{65+} is controlled for (10 and 20 %, respectively), including the TTD^{a} variables increases the expected costs for all three kinds of services, a result driven by the relatively high age coefficients for the oldest individuals[14].

[14]The pattern is similar if we use sex-specific estimates (Appendix 2.A7/2.A8).

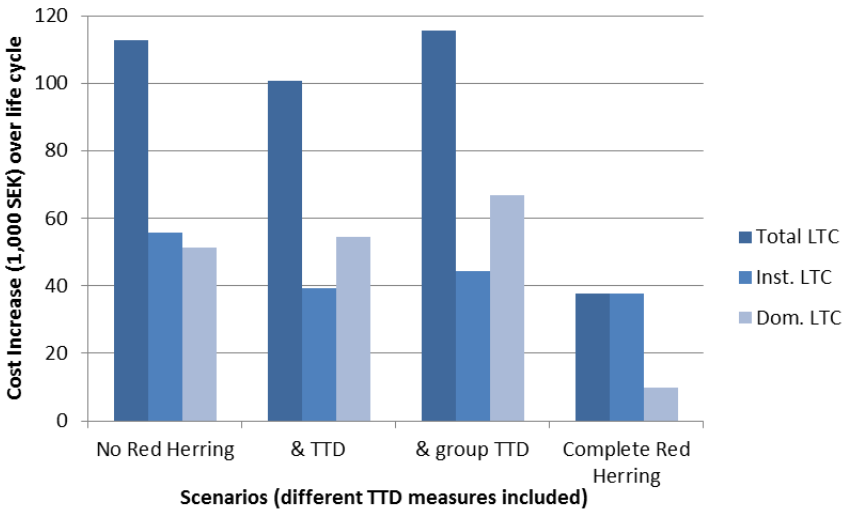


Figure 2.1 Total cost increase per capita associated with an increase in life expectancy of one year (various scenarios)

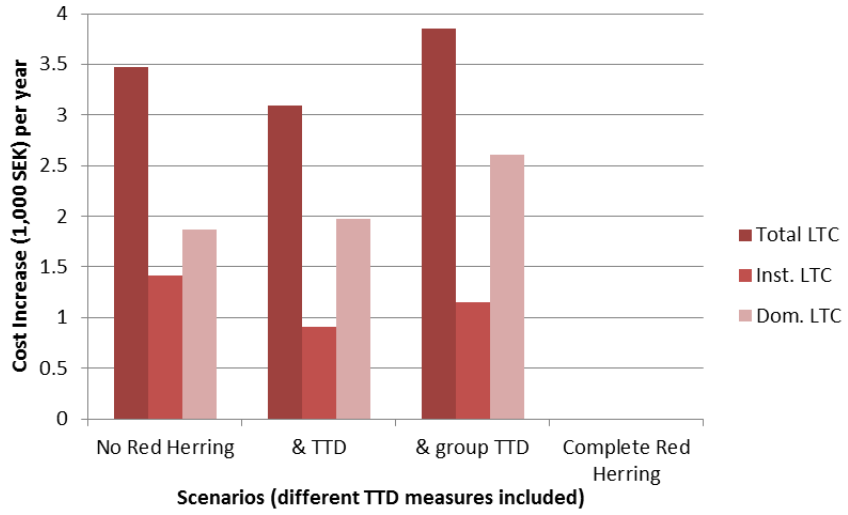


Figure 2.2 Total cost increase per capita and per life year associated with an increase in life expectancy of one year (various scenarios)

The projections of total LTC expenditures for the next 100 years are derived using the coefficients from three estimated specifications, s. The projections are based on regressions when age is accounted for, when TTD^{65+} is included and finally when age-specific TTD is included. Since, from a policy perspective, it is very important to determine the extent to which such future expenditures will be sustainable, we calculate the projected total yearly costs in relation to future fiscal potential. In

doing so, we rescale the projected expenditures for the 65+ population, (N_{65}), with N_p, the number of all inhabitants and the 15- to 65-year-old inhabitants. Formally, these projections are written as:

$$P = x_t^{'}\beta_s \times \frac{N_{65}}{N_p}. \tag{2.4}$$

Figure 2.3 and 2.4 show that although the TTD variable is strongly correlated with LTC expenditures, including this variable into the specification only modestly affects cost projections. The shape of yearly projected values is very marginally lower in the projection based on the specification where TTD is included, although the difference increases to approximately 5 % over time. Including age-specific TTD effects does not change this pattern. Assuming the validity of the projected age structure, we find expenditures will increase from SEK 11,000 to SEK 20,000 per capita over the next 100 years, and from SEK 17,000 to SEK 35,000 if costs are divided by the potential workforce (15- to 64-year-olds). The shape of the projections shows that expenditure increases will be highest in the near future (until 2040) and then increase moderately. This suggests that future budget impacts of ageing can be expected to be considerable in the short run. However, we must also bear in mind that projections for more distant points in time are based on higher uncertainty, due to underlying assumptions about demographic changes in Sweden. As we do not have information about the degree of this uncertainty, we do not provide confidence intervals for our estimates.

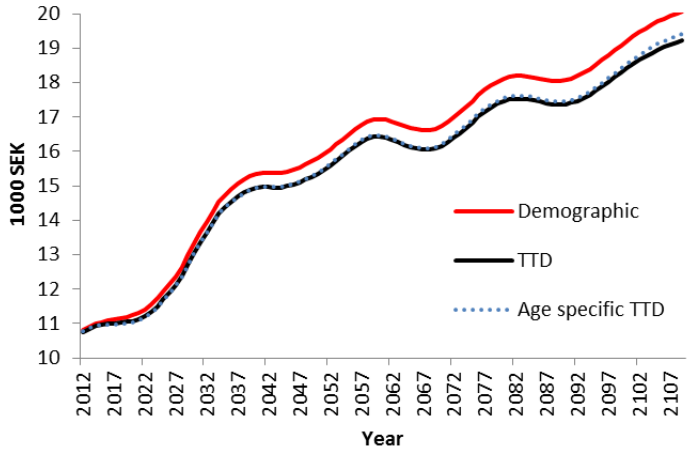


Figure 2.3 Projected LTC expenditures per capita

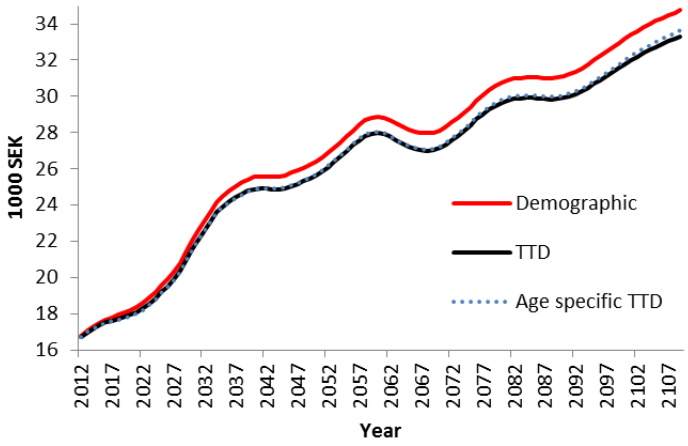


Figure 2.4 Projected LTC expenditures per member of the workforce (15- to 64-year-olds)

2.6 Conclusion

In this chapter, we analysed the effect of ageing and increases in longevity on LTC costs in Sweden. Evaluating these relationships is crucial for policy makers as it allows them to determine future budget expenditures for LTC services.

To increase our knowledge of the 'red herring' hypothesis, we used administrative data on Swedish municipalities, investigating whether TTD is a better predictor of LTC costs than age by controlling for local mortality. One advantage of our study is that we use freely available administrative data that covers all Swedish municipalities that exclusively provide LTC but not healthcare. Compared to other studies using macro data, Sweden provides an excellent framework for our analysis, because the financial redistribution between our units of observation renders budget restrictions for LTC expenditures mostly irrelevant. In addition, we used the panel structure of our data to control for unobserved heterogeneity. The main innovation of this chapter is that our measure for TTD allows us to account for individual end-of-life morbidity effects on the aggregated level more convincingly than existing studies using macro data, by using a retrospective measure for TTD. Based on our empirical findings, we calculated the financial consequences of life expectancy increasing by one year for various scenarios. We also used our estimates to calculate cost projections for the next 100 years based on recent predictions of the Swedish age structure provided by official statistics. It transpires that age is the main determinant of LTC expenditures,

although TTD remains relevant when projecting expenditures.

In addition to our baseline model, we considered several other specifications: Separate estimates for institutional and domiciliary LTC costs revealed a morbidity-related substitution into institutional care in line with evidence based on micro-level analysis. However, another sex-specific specification revealed that the high negative correlation between TTD and domiciliary care is driven mainly by women. This might suggest that, on average, older men do not 'need' to switch to institutional care like women do, because they may still have complementary informal care available to them at the end of their lives courtesy of their spouses. This is a fruitful topic for future research.

Age-specific TTD effects revealed the costs of end-of-life morbidity itself to be strongly related to actual age at death. This issue is usually not accounted for in the literature. Our study found the impact of TTD on LTC costs is mainly driven by a relatively young cohort, the 70- to 74-year-olds. When accounting for heterogeneity in sex as well, this phenomenon is more important for men than women. For men, the effect of TTD lasts even longer: from age 70 to age 79. Whether this finding is a matter of rationing or less demanding care services remains an open question. It would be helpful to assess whether this pattern can be found at the individual level as well.

Our findings show that considerable upward budget shifts due to LTC spending driven by future demographic changes can be expected. Although LTC expenditures for the older population can be explained to a certain extent by TTD (captured by the probability of dying within two years), Sweden's age structure remains more important when focusing on overall LTC. Most importantly, the share of the oldest individuals remains an important determinant of total LTC expenditures. The major significance of ageing is also supported by our cost projections.

However, our study has also some limitations. We are not able to solve the potential endogeneity problem that arises if LTC provision itself influences TTD. However, since – in contrast to healthcare – LTC services mainly emphasise the treatment of chronic illnesses (Norton, 2000), this problem should be of minor relevance in our context. Another possible issue is that we only use data on municipalities that are exclusively responsible for providing LTC expenditures. Although a comparison of observable characteristics suggests our observations are very similar to the situation for the rest of the population, we cannot rule out different outcomes when healthcare is provided at the municipality level. In addition, the time dimension of our data means the population changed to a certain extent, with some individuals leaving the

sample and others entering. This could be a problem if these individuals benefit differently from medical innovations (if not captured by life expectancy, which we account for in our analysis). However, this problem might be more relevant when focusing on curative health care rather than LTC provision. Finally, as we do not include the information in our analysis, our results do not allow any inferences about a LTC receiver's quality of life. However, they highlight that keeping high quality LTC services available will require new ways to use resources efficiently.

Appendix Chapter 2

2.A1 Comparison between municipalities providing and not providing healthcare in 2008

	Healthcare=0		Healthcare=1	
Variable	Mean	Time var.	Mean	Time var.
Total	57.4004	4.5212	57.3170	4.5738
Inst	37.5317	3.9136	35.8417	3.8765
Dom	17.5361	3.7201	18.8287	3.7382
age6569	0.2667	0.0133	0.2622	0.0148
age7074	0.2356	0.0108	0.2311	0.0109
age7579	0.2099	0.0131	0.2085	0.0136
age8084	0.1581	0.0096	0.1605	0.0091
age8589	0.0894	0.0073	0.0932	0.006
age9094	0.0331	0.0041	0.0362	0.0038
age95100	0.0072	0.0016	0.0084	0.0018
mrt	0.0517	0.0047	0.0512	0.0046
mrtL1	0.0536	0.0049	0.0531	0.0048
TTD65	0.1025	0.0066	0.1016	0.0067
TTD6569	0.0076	0.0015	0.0072	0.0015
TTD7074	0.0112	0.0021	0.0105	0.0021
TTD7579	0.0175	0.0033	0.0164	0.003
TTD8084	0.0238	0.003	0.0232	0.003
TTD8589	0.0235	0.0028	0.0239	0.0027
TTD9094	0.0141	0.0023	0.0149	0.0022
TTD95100	0.0045	0.0012	0.0053	0.0014
medin 08	156.5511	10.63	150.368	11.1718
privcare	0.0612	0.0477	0.0412	0.0476
wom65	0.5575	0.0063	0.5518	0.006
density	157.7674	17.0796	92.0616	6.6494
rightwing	0.3369	0.0233	0.3538	0.0222
taxrate	0.2138	0.003	0.2122	0.0045
mrtl1	0.0519	0.0047	0.0516	0.0046
lifexp	18.4242	0.8155	18.7221	0.7898
N	1,589		1,419	

Mean comparison of municipalities by healthcare provision. Healthcare=0 represents the sample used in our analysis, whereas Healthcare=1 represents excluded observations. Time var. represents standard deviation *within* time to show that changes over time are similar for both groups.

2.A2 Variable definitions

Variable	Description
Total	total LTC costs per 65+ inhabitant in 1,000 SEK
Inst	institutional LTC costs per 65+ inhabitant in 1,000 SEK
Dom	domiciliary LTC costs per 65+ inhabitant in 1,000 SEK
age6569	proportion of the 65+ population aged 65-69
age7074	proportion of the 65+ population aged 70-74
age7579	proportion of the 65+ population aged 75-79
age8084	proportion of the 65+ population aged 80-84
age8589	proportion of the 65+ population aged 85-89
age9094	proportion of the 65+ population aged 90-94
age95100	proportion of the 65+ population aged 95 to 100+
mrt	mortality rate among the 65+ population
mrtL1	Next year's mortality rate among the 65+ population
TTD65+	two-year mortality rate for the 65+ population
TTD6569	two-year mortality rate for the population aged 65-69
TTD7074	two-year mortality rate for the population aged 70-74
TTD7579	two-year mortality rate for the population aged 75-79
TTD8084	two-year mortality rate for the population aged 80-84
TTD8589	two-year mortality rate for the population aged 85-89
TTD9094	two-year mortality rate for the population aged 90-94
TTD95100+	two-year mortality rate for the population 95 to 100+
medinc65_08	median income of 65+ population
privcare	share of people using private LTC services
wom65	share of women (65+)
density	population density
rightwing	share of seats occupied by right-wing parties
taxrate	local tax rate
mrtl1	last period's mortality rate among the 65+ population
lifexp	life expectancy for members of the population aged 65

2.A3 Sex-specific summary statistics

Variable	Mean	Std. Dev.	Min.	Max.
age6569m	0.131	0.02	0.089	0.236
age7074m	0.111	0.01	0.081	0.144
age7579m	0.093	0.009	0.066	0.125
age8084m	0.064	0.009	0.024	0.099
age8589m	0.032	0.006	0.008	0.053
age9094m	0.01	0.003	0.003	0.022
age95100m	0.002	0.001	0	0.008
lifexpm	16.838	1.321	13.189	22.454
TTD65m	0.048	0.007	0.028	0.073
TTD6569m	0.005	0.002	0	0.016
TTD7074m	0.007	0.002	0	0.017
TTD7579m	0.01	0.003	0.003	0.023
TTD8084m	0.012	0.003	0.003	0.025
TTD8589m	0.01	0.003	0.001	0.021
TTD9094m	0.005	0.002	0	0.014
TTD95100m	0.001	0.001	0	0.005
age6569w	0.136	0.016	0.097	0.221
age7074w	0.125	0.01	0.093	0.153
age7579w	0.117	0.012	0.071	0.165
age8084w	0.094	0.012	0.043	0.138
age8589w	0.058	0.01	0.029	0.09
age9094w	0.023	0.005	0.007	0.042
age95100w	0.006	0.002	0	0.014
lifexpw	19.994	1.229	14.057	25.222
TTD65w	0.052	0.007	0.031	0.076
TTD6569w	0.003	0.001	0	0.007
TTD7074w	0.004	0.001	0	0.011
TTD7579w	0.008	0.002	0.001	0.017
TTD8084w	0.012	0.003	0.003	0.023
TTD8589w	0.014	0.003	0.004	0.029
TTD9094w	0.009	0.002	0.002	0.018
TTD95100w	0.003	0.001	0	0.01

Summary statistics for the explanatory variables used for sex-specific estimates. w(omen) and m(en) indicate which sex each variable corresponds to.

2.A4 Sex-specific institutional LTC expenditures

	(1) Inst	(2) Inst	(3) Inst	(4) Inst	(5) Inst
				Men	Women
age6569m	-6.34	-6.38	-4.74	-4.89	
	(30.09)	(29.69)	(30.11)	(29.67)	
age7074m	7.52	1.40	3.13	-8.24	
	(35.06)	(35.90)	(36.03)	(36.35)	
age7579m	-28.81	-37.39	-34.99	-43.89	
	(41.94)	(41.36)	(41.99)	(43.94)	
age8084m	-5.81	-17.32	-14.19	-18.01	
	(49.19)	(51.01)	(51.26)	(53.71)	
age8589m	-15.75	-34.60	-30.26	7.44	
	(70.40)	(75.07)	(75.05)	(80.50)	
age9094m	90.98	58.29	63.65	150.57	
	(110.29)	(116.90)	(116.56)	(143.91)	
age95100m	-120.94	-96.52	-87.78	-150.26	
	(227.01)	(218.12)	(220.83)	(358.92)	
age6569w	53.12	53.60*	55.27*		50.88
	(32.67)	(32.07)	(32.31)		(32.33)
age7074w	53.46	51.89	53.62		49.76
	(33.04)	(33.00)	(32.86)		(33.80)
age7579w	11.60	2.63	4.52		14.71
	(32.46)	(32.84)	(32.90)		(34.11)
age8084w	75.96*	65.33	67.76*		74.20*
	(39.92)	(40.82)	(40.86)		(43.86)
age8589w	93.57**	70.73*	73.28*		60.14
	(38.58)	(39.76)	(39.67)		(42.22)
age9094w	131.46*	88.42	90.96		77.41
	(75.65)	(78.43)	(78.01)		(87.08)
age95100w	215.05	153.87	158.04		200.92
	(153.52)	(156.87)	(157.61)		(203.97)
TTD65m		65.46*	58.14		
		(38.14)	(41.59)		
TTD65w		109.98***	103.99***		
		(34.67)	(37.43)		
lifexpm			-0.04	-0.02	
			(0.12)	(0.12)	
lifexpw			-0.04		-0.03
			(0.10)		(0.10)
TTD6569s				67.01	214.90*
				(92.84)	(127.68)
TTD7074s				148.50*	247.47**
				(84.30)	(100.15)
TTD7579s				102.49	1.78
				(74.34)	(75.00)
TTD8084s				88.86	56.68
				(77.74)	(71.85)
TTD8589s				-52.75	141.72**
				(79.28)	(69.47)
TTD9094s				-112.95	120.44
				(122.73)	(85.61)
TTD95100s				207.60	1.45
				(321.43)	(178.44)
Observations	1,589	1,589	1,589	1,589	

Fixed effects estimates for institutional LTC expenditures when age-specific control variables are included. We included sex-specific coefficients for the 65-69 age group as we do not rely on a constant in this specification. Column 1 accounts for age, column 2 includes TTD and column 3 uses life expectancy. In columns 4 and 5 (coefficients taken from a single regression), we allow for age- and sex-specific TTD effects on institutional LTC expenditures. Year dummies are included and the regressions are weighted by the square root of each municipality's average 65+ population. The unit of observation is a Swedish municipality for the period 1998-2008. Standard errors are clustered at the municipality level. * p<0.10, ** p<0.05, *** p<0.01

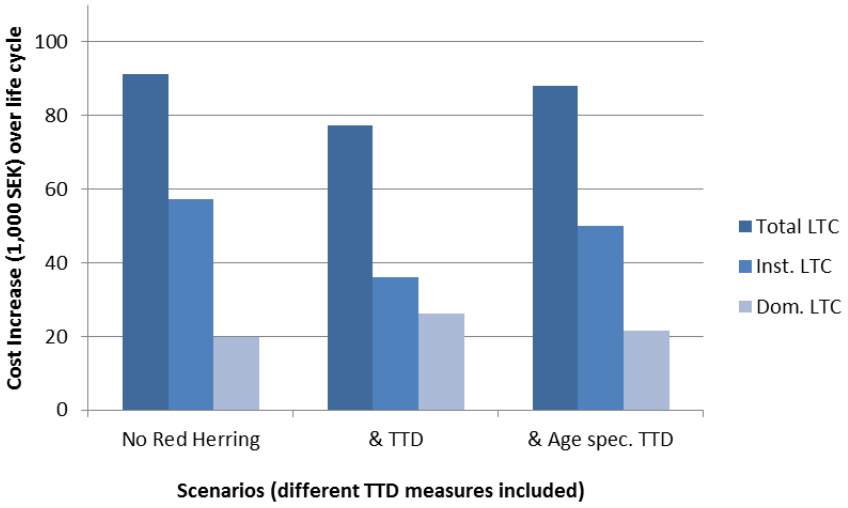
2.A5 Sex-specific domiciliary LTC expenditures

	(1) Dom	(2) Dom	(3) Dom	(4) Dom	(5) Dom
				Men	Women
age6569m	33.24	30.50	31.06	30.91	
	(23.33)	(23.43)	(23.77)	(23.52)	
age7074m	24.22	19.63	20.06	16.17	
	(27.69)	(27.55)	(27.56)	(28.18)	
age7579m	18.66	12.20	12.82	10.73	
	(30.00)	(30.25)	(31.07)	(32.53)	
age8084m	75.85*	64.60	65.17	71.23	
	(39.90)	(41.88)	(42.73)	(43.28)	
age8589m	103.42*	86.87	87.62	87.00	
	(55.06)	(60.83)	(60.29)	(64.36)	
age9094m	152.84	130.56	131.38	125.74	
	(93.44)	(97.12)	(97.09)	(121.78)	
age95100m	325.83*	286.53	287.77	408.56	
	(175.56)	(176.17)	(177.09)	(272.44)	
age6569w	-15.37	-11.42	-10.74		-7.44
	(22.54)	(22.05)	(22.03)		(22.76)
age7074w	-19.76	-15.22	-14.42		-17.11
	(27.25)	(27.30)	(27.02)		(26.97)
age7579w	18.41	27.08	28.04		17.30
	(26.75)	(27.64)	(27.61)		(27.61)
age8084w	-14.83	-4.02	-2.87		-17.15
	(31.71)	(32.64)	(33.05)		(35.03)
age8589w	49.84	66.27*	67.76*		73.10*
	(33.01)	(35.42)	(35.61)		(37.52)
age9094w	104.96*	129.74**	131.65**		183.78***
	(58.75)	(61.77)	(61.46)		(66.52)
age95100w	146.10	183.81	186.78		233.29
	(114.14)	(117.52)	(118.41)		(154.54)
TTD65m		19.35	19.24		
		(30.20)	(33.30)		
TTD65w		-59.04**	-63.55*		
		(28.50)	(32.73)		
lifexpm			-0.00	0.02	
			(0.09)	(0.09)	
lifexpw			-0.03		-0.01
			(0.08)		(0.08)
TTD6569s				1.91	-178.11*
				(90.29)	(106.98)
TTD7074s				78.05	-40.94
				(69.33)	(83.05)
TTD7579s				33.85	52.07
				(56.42)	(58.54)
TTD8084s				6.56	30.42
				(60.78)	(62.61)
TTD8589s				19.68	-96.55*
				(63.75)	(50.42)
TTD9094s				46.69	-208.08***
				(94.60)	(75.19)
TTD95100s				-158.21	-98.52
				(254.08)	(137.84)
Observations	1,589	1,589	1,589	1,589	

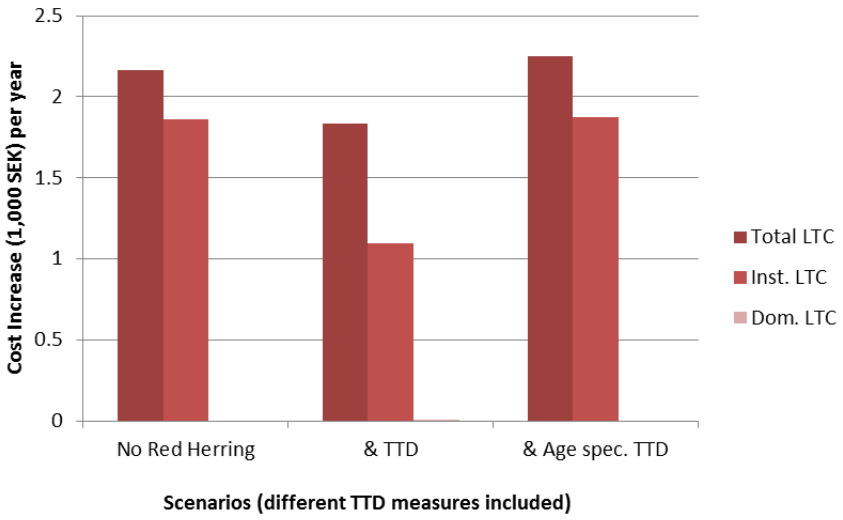
Fixed effects estimates for domiciliary LTC expenditures when age-specific control variables are included. We included sex-specific coefficients for the 65-69 age group as we do not rely on a constant in this specification. Column 1 accounts for age, column 2 includes TTD and column 3 uses life expectancy. In columns 4 and 5 (coefficients taken from a single regression), we allow for age- and sex-specific TTD effects on domiciliary LTC expenditures. Year dummies are included and the regressions are weighted by the square root of each municipality's average 65+ population. The unit of observation is a Swedish municipality for the period 1998-2008. Standard errors are clustered at the municipality level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

2.A6: Changes in cost increase per 65+ life year

$$\triangle LTC_{it} = \begin{cases} \sum_{k=65}^{\infty} \left(\alpha_j + \beta_j^k\right) \left(\frac{s_{2008}^{k-1}}{\sum_{k=65}^{\infty} s_{2008}^{k-1}} - \frac{s_{2008}^{k}}{\sum_{k=65}^{\infty} s_{2008}^{k}} \right) & with\ j = i \quad , for\ i = 1, 2 \\ \sum_{k=65}^{\infty} \left(\left(\alpha_j + \beta_j^k\right) \left(\frac{s_{2008}^{k-1}}{\sum_{k=65}^{\infty} s_{2008}^{k-1}} - \frac{s_{2008}^{k}}{\sum_{k=65}^{\infty} s_{2008}^{k}} \right) + \delta_k^j \left(\frac{TTD_{2008}^{k-1}}{\sum_{k=65}^{\infty} s_{2008}^{k-1}} - \frac{TTD_{2008}^{k}}{\sum_{k=65}^{\infty} s_{2008}^{k}} \right) \right) & with\ j = i \quad , for\ i = 3 \\ \frac{\alpha_j + \sum_{k=65}^{\infty} \left(\alpha_j + \beta_j^k\right) s_{2008}^{k-1}}{\sum_{k=65}^{\infty} s_{2008}^{k-1}} - \frac{\sum_{k=65}^{\infty} \left(\alpha_j + \beta_j^k\right) s_{2008}^{k}}{\sum_{k=65}^{\infty} s_{2008}^{k}} & with\ j = 1 \quad , for\ i = 4 \end{cases}$$



2.A7 Total cost increase per capita associated with an increase in life expectancy of one year (various scenarios based on sex-specific estimates)



2.A8 Total cost increase per capita and per life year associated with an increase in life expectancy of one year (various scenarios based on sex-specific estimates)

Chapter 3

The Ability to Memorise and Participation in the English Bowel Cancer Screening Programme

3.1 Introduction

Classical economic theory suggests that a rational individual will maximise their expected utility and base their decisions on the expected outcomes in different scenarios. In a situation governed by standard economic assumptions, where an individual is fully informed and can use this information in a decision making process without further restrictions, such an analysis is perfectly reasonable. However, individuals often face restrictions in the transmission of information when making decisions (Kahneman, 2003). These restrictions may induce additional costs or prevent an individual from making proper decisions if pertinent information cannot be fully used. One example of such a constraint is cognitive ability, which may be a strong barrier to making economic or health related decisions. This is especially relevant for the elderly, as cognitive decline has been identified as part of the human life cycle (Salthouse, 2009), potentially distorting corresponding health decisions in the elderly.

Many European countries have begun to involve patients in the medical decision making process by implementing preventive cancer screening programmes – mostly focused on breast cancer – which may help contain costs and reduce healthcare expenditures (Maciosek et al., 2006) in times of substantial fiscal burden. The UK programme for breast cancer screening is very successful and has been thoroughly evaluated (e.g. Marmot et al., 2013). However, there is little evidence about partic-

ipation in the bowel cancer screening programme launched 2006 in England. The gerontological literature suggests that a decline in cognitive abilities is a general phenomenon throughout the life cycle, especially in the elderly population (Salthouse, 2009). As the target population of the bowel cancer programme is individuals aged 60+, we investigate whether memorisation ability is a determinant of an individual's bowel cancer screening decision. The main contribution of this chapter is our assessment of the link between memorisation ability and bowel cancer screening from a causal perspective by mitigating common identification problems such as reverse causality and omitted variable bias. We use longitudinal data from the ELSA project, a national representative survey from England, which has the advantage that we can use objective information about cognitive skills based on respondents' performance on specific tasks. We use individual changes in the ability to memorise information over time as our treatment to estimate the impact of declining memory on the likelihood of participating in a public bowel cancer screening programme. Our identification strategy is based on covariate adjustment by applying linear probability models, propensity score matching and a double robust estimator. We find that a decline in memorisation ability is negatively related to an individual's participation in bowel cancer screening.

The chapter is structured as follows: The next section provides a literature overview that discusses existing evidence from an economic and psychological/medical perspective. We then provide an overview of the NHS bowel cancer programme, followed by a description of our identification strategy and the data. After presenting descriptive statistics and our empirical findings, we discuss our results, followed by a conclusion that emphasises policy implications and potential future research.

3.2 Literature

The economics literature is home to a huge debate about factors that explain so-called 'healthy behaviours'. Cutler and Glaeser (2005) highlight the huge uncertainty surrounding possible mechanisms, as both levels and differences in health behaviours such as smoking or preventive activity are not very strongly correlated. There is some economic evidence for cognition and health behaviour which emphasises the importance of education and complex health decisions due to quicker adoption of new technologies (Glied and Lleras-Muney, 2008) or higher intellectual curiosity (Cutler and Lleras-Muney, 2010). However, despite a consensus about the empirical asso-

ciation between education and health, the causal pathways are difficult to discern, due to the problem of reverse causality. Many empirical studies focus on the impact of schooling on health outcomes (Eide and Showalter, 2011) and many distinguish between direct and indirect effects of education. Lange (2011), for example, analyses the role of information in the impact of education on breast cancer screening using data from the US. Mocan and Altindag (2014) also use US data to identify the role of education in health production. However, as the empirical evidence from these studies is inconclusive, it is natural to directly assess the impact of cognitive abilities on health behaviours such as preventive healthcare activities.

Although there is evidence for the direct association between cognitive abilities and health outcomes later in life (Kaestner, 2009), economic literature on the relationship between cognitive abilities and preventive medical activities remains scarce. Avitabile et al. (2011) assess the extent to which education and cognitive abilities such as cognitive fluency and numeracy complement or supplement health promotion programmes in European countries.Wübker (2012, 2014) uses verbal fluency as a proxy for cognition and analyses its impact on mammography take-up across Europe. Wübker (2012) also uses the ability to memorise (among other variables) as a driver of an individual's decision to participate in mammography screening. Here, memory is not found to be relevant across the estimated specifications. This is surprising, because the psychological literature emphasises the potential role of memorisation on medical decision making, as discussed in the following.

The role of cognitive abilities in health behaviours is discussed in more depth in the psychological and epidemiological literature. One major branch of this literature emphasises medication adherence. Hayes et al. (2009) find that people with a relatively high cognitive function take their medication more regularly than a control group. The authors use both a very general measure for cognition and a measure for the ability to memorise. Insel et al. (2013) emphasise the role of working memory and executive function to evaluate whether anchoring strategies can improve medication adherence in the elderly. The results of their study are not yet available.

A relatively large research area in the psychological literature focuses on the importance of health literacy, which also involves cognitive processing. A recent review by Oldach and Katz (2014) reveals mixed results in the association between health literacy and cancer screening participation. Recent medical evidence on colorectal cancer screening in England (Kobayashi et al., 2014) suggests that health literacy may be a driver of cancer screening participation in the elderly. One drawback of this study is that the findings merely represent statistical associations, since their

analysis does not use other important factors such as measures for cognitive abilities. Recent evidence suggests that the ability to memorise is an important driver of health literacy itself (Wilson et al., 2010), potentially explaining why some older people participated in the NHS bowel cancer programme while others did not.

The lack of evidence for the relationship between memorisation ability and cancer screening is evident in England, where the bowel cancer programme targets people aged between 60 and 70 years (currently until age 75). It is well established that memorisation ability declines with age, and can even accelerate in old age (Salthouse, 2009), potentially becoming an obstacle to medical decision making. Most existing studies merely reflect descriptive patterns rather than causal relationships, since it is not possible for researchers to randomly assign cognitive ability to participants of a study. We contribute to the literature by applying covariate adjustment methods to reveal a potential causal relationship between an individual's memorisation ability and their participation in the English colorectal cancer screening programme.

3.3 The English Bowel Cancer Programme and Memorisation Ability

There are two steps to the screening process. First, an invitation letter is sent to eligible candidates, men and women aged 60 to 70. It contains information about the benefits of screening. A week later, respondents are sent a cancer screening kit, which instructs them to collect samples and return them for laboratory testing[1]. There are several reasons why cognitive distortions may affect participation in screening. First, the costs and benefits outlined on the information leaflet must be understood and memorised to proceed to the second step. If a respondent does not fully remember the contents of the information leaflet, they are less likely to participate in the actual screening. Second, a decline in cognition can make participating in cancer screening more costly, as decisions cannot be implemented as efficiently as before, e.g., more time and effort is required to carry out everyday decisions.

[1]Further details on the programme procedures can be found at: http://www.cancerscreening.nhs.uk/bowel/index.html

3.4 Analysis

3.4.1 Empirical Approach

We use the (negative) change in someone's ability to memorise between two waves as our treatment because we assume that an analysis based on levels is more likely to be confounded by other factors (cf. Frey, 1990). More specifically, we choose the difference $T = M_{it-1} - M_{it}$ as our treatment, where t is a survey wave indicator and M represents an individual's level of memory. A similar strategy is used by Decker and Schmitz (2015), who analyse the impact of an individual health shock on risk preferences. We assume that, given characteristics X, a decline in memorisation ability, T, is exogenous in the short run of the potential outcomes Y_1 and Y_0, which represent participation in the bowel cancer programme when being 'treated' and 'not treated', respectively. More formally, we assume $Y_1, Y_0 \perp T|X$. Although we use the panel structure of our data to model the causal relationship between our variables using data from several waves, our estimates rely solely on cross-sectional variation. More detailed information about the estimated specifications can be found in the next section.

We begin by assessing the impact of a change in memorisation ability on future screening participation by applying a linear probability model (LPM). We provide both a bivariate regression and a specification where other variables (from the previous period) such as lagged health information and sociodemographic characteristics are included into the regression model, mainly to increase the precision of our estimates but also to control for potential confounders. One important missing factor may be unobserved preferences that determine an individual's decisions, leading to changes in both memorisation ability and screening behaviour. To mitigate the potential problem that such preferences (if not already captured by the observable health information) are correlated with both changes in cognition and an individual's decision to participate in cancer screening, we control for further observable characteristics which we interpret as indicators of 'healthy behaviour', social activities and other characteristics.[2] The corresponding estimates are based on the following equation:

$$y_i = \beta_0 + \beta_1 T_i + X^{'}\beta + \epsilon_i \tag{3.1}$$

Second, we apply propensity score matching by explicitly modelling the treatment

[2] A more detailed discussion of the variables used is provided below.

'decline in memorisation ability' and then calculating the treatment effect using an individual's probability of 'being treated'. To model the treatment decision, we estimate a logit model where $\mathbb{1}(T > 0)$ (i.e., a variable that takes a value of 1 if there is a measurable decline and zero otherwise) serves as our dependent variable.

The matching estimator has the advantage that it reduces the potential problem of insufficient common support while introducing relatively few parametric assumptions (Lechner, 2009).

Third, we apply a double robust estimator (Robins et al., 1995), a weighted regression approach that allows us to deconfound both the outcome and treatment equations and provides unbiased estimates even if one of both equations is wrongly specified. Following Morgan and Todd (2008), we assign a weight of 1 to all treated observations and a weight of $\frac{\hat{p}_i}{1-\hat{p}_i}$ to the control units to estimate the average treatment effect on the treated (ATT). The estimated probabilities, $\hat{p}_i$, are used to turn the control group into a sample that is representative of the treatment group.

3.4.2 Data and Specification

To capture the impact of a decline in memorisation ability on participation in colorectal cancer screening in England, we use information from the 3rd, 4th, 5th and 6th wave of the ELSA English household survey, which were conducted between 2006 and 2013. There is a gap of around two years between the interview dates. We restrict the analysis to individuals who participated in all four ELSA waves. This data source strongly complements our research question, as it is representative for the population aged 50+ and contains important health-related information. One main variable of interest captures a change in memorisation ability. ELSA respondents participated in several tests to measure their performance in different areas of cognitive ability. Several ELSA waves contained tests on memorising information. We aggregate information from three tasks to derive a summary measure of respondents' memorisation ability. The first two tasks measure respondents' ability to memorise a list of 10 words (directly for the first task, and after some delay for the second). The third task measures respondents' ability to memorise the components of the current date. The aggregated variable used in our analysis varies between 0 and 24 points. Other research has shown that aggregated variables representing cognitive abilities derived from ELSA can be used very effectively in longitudinal comparisons (Steel et al., 2004). When applying the LPM, we chose the decline in cognitive abilities between waves three and four for our treatment variable. We chose this timeframe because our outcome variable is only observable after this point. To create the propensity

scores used in our matching approach and the double robust estimator, we use this difference to create a dummy representing the binary treatment, as discussed in the previous section.

Our dependent variable captures the decision to use a bowel screening kit. This variable is available in and derived from ELSA waves 5 and 6. During these two waves, individuals were asked whether they had participated in the cancer screening programme and how recently they last participated. As we want to correctly model the decision making process over time with our empirical specification (to rule out reverse causality), we assigned the fourth wave as the point when our artificial treatment occurs. Our dependent variable has a value of 1 if the respondent participated in screening after the wave 4 interview and 0 otherwise[3]. To ensure a clean control group, we exclude all individuals who participated in bowel cancer screening before the interview date during wave 4. This information is retrospectively provided from wave 5 onwards. The sample is restricted to people aged 60 to 67, because we know this group is eligible for screening and was invited via information leaflets. As we evaluate the probability of screening within a time range of 4 years after the year of treatment, this age range guarantees that all individuals in our sample were actually invited for screening during the 4 years after treatment. Hence, we rule out the possibility that our results are affected by heterogeneity (in information) due to the fact that older individuals were no longer invited for screening.

We include control variables in our empirical specifications to account for potential confounders which may be associated with both a decline in memorisation ability and participation in cancer screening. The time-variant control variables are taken from wave 3 to prevent them being affected by our treatment variable, which would bias our estimate. Figure 3.1 shows the timeframe we use for our identification strategy.

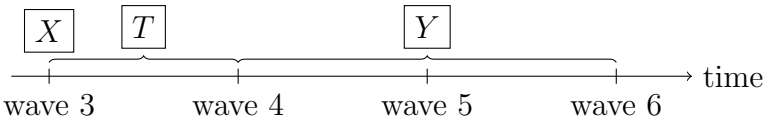


Figure 3.1 Definition of outcome, treatment and control variables within the chosen timeframe

Table 3.1 shows all the variables used in our analysis. The most important control

[3]Due to a substantial number of missing values for colorectal cancer screening in wave 5, we are unfortunately unable to distinguish between short- and long-term effects

variables represent an individual's health status, captured by self-assessed health (where 1 indicates bad/very bad health and 4 very good health), six disease-specific illness categories and problems with everyday activities. Of the six illness variables, the most important is 'ill5', which captures mental disorders such as dementia or Alzheimer's. In addition, ex ante differences in activities of daily living (ADL) are of special interest to our analysis, as the programme targets older people and problems with ADL usually increase with age (Gremeaux et al., 2012).

Table 3.1 Descriptive statistics

Variables	Total mean	Min	Max	Mean (T=1)	Mean (T=0)
screening	0.814	0	1	0.795	0.829
memory_decl	0.137	-14	11	2.863	-2.032
memory	15.411	4	24	16.755	14.342
educ_low	0.190	0	1	0.175	0.202
educ_mid	0.584	0	1	0.586	0.582
educ_high	0.226	0	1	0.240	0.216
wealth	382.012	-42.051	20818.01	345.154	411.337
adl	0.190	0	1	0.188	0.192
health_1	0.052	0	1	0.049	0.054
health_2	0.207	0	1	0.198	0.215
health_3	0.461	0	1	0.470	0.454
health_4	0.280	0	1	0.283	0.278
ill1	0.506	0	1	0.494	0.517
ill2	0.166	0	1	0.163	0.169
ill3	0.396	0	1	0.400	0.393
ill4	0.072	0	1	0.069	0.075
ill5	0.159	0	1	0.157	0.160
ill6	0.130	0	1	0.136	0.126
female	0.547	0	1	0.525	0.565
age	63.118	60	67	63.247	63.016
black	0.015	0	1	0.017	0.013
couple	0.789	0	1	0.792	0.786
smoker	0.143	0	1	0.147	0.139
smoke_past	0.481	0	1	0.501	0.466
organisation	0.465	0	1	0.475	0.457
phys_act2	0.166	0	1	0.152	0.177
phys_act3	0.559	0	1	0.584	0.538
phys_act4	0.254	0	1	0.242	0.263
life_exp	49.256	0	100	49.208	49.295
N	1564			693	871

Our analysis also uses participants' self-assessed probability of living to 85 to capture heterogeneity not already captured by the other controls for health. This

variable may also capture some heterogeneous preferences for health consumption in the spirit of Fang et al. (2007). Another control variable is the baseline ability to memorise, because any change in cognitive abilities may also depend on initial levels. Inclination towards healthy behaviour and social activities are captured by an individual's smoking status, smoking history and whether the individual is active in more than one organisation. In addition, we allow for differences in physical activity levels. Other common sociodemographic characteristics such as sex, age, education and household wealth are also accounted for.

Comparing the last two columns in Table 3.1 (representing our binary treatment) show that most variables are very similar across the treatment and control group. The share of women is slightly lower in the treatment than in the control group (53 compared to 57 %). The share of former smokers is also higher in the treatment than in the control group. Substantial differences were found in baseline ability to memorise, where the treatment group was able to answer an average of around two additional questions correctly compared to the control group. Average (financial) wealth was lower in the treatment group than the control group. There was also some heterogeneity in physical activity levels between both groups.

3.4.3 Findings

The results in Table 3.2 show the estimates for the LPM. We see that a reduction in memorisation ability is a negative predictor of future screening participation (Column 1). Reducing the test score by one point (over time) decreases participation by around 1 %. This finding is strongly statistically significant. If other covariates are included, the effect is slightly higher (Column 2). The signs of most covariates are in line with our expectations, although not all are significant. Being female, lagged memorisation ability, wealth, general health status, being in a relationship and being a member of organisations are all positively associated with screening behaviour. The dummy for high education has a negative point estimate, which may initially seem odd, but both its economic and statistical relevance are very low. Smoking behaviour, also an indicator of unobserved preferences for healthy living, is negatively related to screening participation. As expected, level of physical activity correlates positively with bowel cancer screening.

We now use our matching estimates[4] to evaluate our artificial binary treatment. The main difference from the regression-based approach – which assumed all poten-

[4]We use the statistics package psmatch2 (Leuven and Sianesi, 2014).

Table 3.2 Impact of a cognitive decline on screening (OLS)

	Specification 1		Specification 2	
Variables	Coef.	t-stat	Coef.	t-stat
memory_decl	-0.009***	(-2.82)	-0.012***	(-3.41)
memory			0.007*	(1.82)
educ_mid			0.021	(0.78)
educ_high			-0.019	(-0.56)
wealth			0.0000272**	(2.29)
adl			0.037	(1.28)
health_2			0.181***	(3.18)
health_3			0.181***	(3.17)
health_4			0.155**	(2.57)
ill1			-0.033*	(-1.72)
ill2			-0.010	(-0.35)
ill3			-0.010	(-0.49)
ill4			-0.069*	(-1.84)
ill5			0.013	(0.46)
ill6			0.012	(0.42)
female			0.051**	(2.50)
black			0.000	(0.00)
couple			0.053**	(2.01)
smoke			-0.157***	(-4.43)
smoke_past			-0.006	(-0.31)
phys_act2			0.214**	(2.47)
phys_act3			0.213**	(2.50)
phys_act4			0.227***	(2.61)
organisation			0.058***	(2.84)
life_exp			0.000	(0.41)
cons	0.815***	(79.60)	0.273**	(2.34)
N	1564		1564	
$adj.R^2$	0.01		0.07	

Notes: Age fixed effects included in specification 2. Standard errors are clustered at the household level. t-statistics (in parentheses) * p<0.10, ** p<0.05, *** p<0.01.

tial confounders in the outcome equation were considered – is the need to include all potential confounders related to the 'treatment assignment'. For the matching algorithm, we employ radius matching with a calliper of 0.01 and excluded 14 observations from the matching procedure due to insufficient overlap at the right hand tail of the distributions of the propensity scores between the treatment and control group. Applying the PS-matching estimator again shows a strong negative impact of reduced memorisation ability on screening participation (Treatment effect: -0.052; t-stat: -2.04). In a logit regression, where all variables are considered to estimate the propensity score[5], wealth and being female are negatively associated with a decline in cognitive abilities, whereas smoking has a positive impact. We also find a positive impact of lagged memory as expected, since very high values before our treatment year will have a tendency to swing back towards the population mean (Yudkin and Stratton, 1996). In addition, conditioning on the 'lagged treatment' captures baseline differences in memorisation ability between the treatment and control group.

In applying the double robust estimator, a weighted regression which combines the covariate adjustment strategies from pure linear regression and propensity score matching, we follow Crump et al. (2009) and drop observations from our analysis that have estimated propensity scores < 0.1 or > 0.9, to compensate for insufficient overlap. These estimates also suggest a causal effect between decline in memorisation ability and future screening (Treatment effect: -0.045; t-stat: -2.03). Although slightly smaller than the estimate from propensity score matching, we conclude from this result that the estimated effect is robust to misspecification, assuming that one model is correctly specified.

To shed some light on how reduced memorisation ability can affect screening participation, we further utilise the fact that our measure for memorisation is based on different tasks respondents had to perform during the interview. Hence, we separately estimate the impact of decreased performance in each of these tasks on future bowel cancer participation.

The resulting estimates, provided in Table 3.3, show the ability to memorise with some delay is particularly important in the context of the decision to undergo colorectal cancer screening, while the ability to remember components of the date is the least important. This suggests that the main trigger is the capacity to memorise and numerical skills, which are another important dimension of cognitive ability, are less important. We are not able to assess this issue in more detail, as the information on numeracy was not provided in the ELSE waves we are interested in.

[5]See Appendix 3.A1.

Table 3.3 Treatment heterogeneity

Treatment	Coef	z-stat	N
decline_total	-0.045	(-2.03)	1466
decline_date	-0.038	(-1.08)	1032
decline_words	-0.019	(-0.83)	1489
decline_delay	-0.042	(-1.96)	1498

Notes: Outcomes for the double robust estimator. Observations with propensity scores below 0.1 and above 0.9 are excluded. The resulting coefficients show the treatment effects when using a decline in different aspects of memorisation as the treatment. Standard errors are clustered at the household level. All control variables are accounted for in underlying estimates.

Although our findings are closely related to the literature on medication adherence, the relationship between an individual's memory and the ability to use medication on a regular basis is quite different from participation in a cancer screening programme. Whereas in the former case the decision about willingness to take a specific drug is already made, memory may have an additional impact on this decision in the latter case. This may happen if an individual cannot memorise the information about the programme between the time they receive information about it and later receipt or use of the screening kit. Evaluating the specific mechanism which leads to a reduction in screening participation is an important question for future research.

Our results show the negative impact of a decrease in memorisation ability on participating in the programme for bowel cancer screening. This shows the importance of considering specific characteristics of a programme's target population when designing health economic policies in order to more effectively increase public health or decrease healthcare expenditures.

3.5 Conclusion

This chapter analysed the causal impact of the ability to memorise information on participation in a national bowel cancer screening programme in England. After discussing existing evidence on this research question from different scientific fields, we introduced our empirical strategy and the data. Our main contribution is that we used four waves from the ELSA survey, a national representative data for the English elderly population, reducing common problems such as omitted variable bias and reverse causality. We provided robust evidence about the suggested relationship and

implemented regression, propensity score matching and double robust techniques in our empirical strategy. When distinguishing between different dimensions of memory, we find evidence that the estimated effect is driven to a large extent by the ability to memorise information with some delay.

Our findings are related to research into medication adherence, which has found wider consideration in the medical and psychological literature (e.g. Hayes et al., 2009; Insel et al., 2013). Studies have shown that the ability to memorise is a fundamental issue in the successful treatment of several illnesses via medication. Our findings contribute to this, as we show that memorisation ability can also affect health via the channel of preventive medical behaviour. This suggests that heterogeneity in mental ability along further dimensions of a target population must be considered when optimising the planning and supply of healthcare provision. Another finding is that the main driver relating an individual's memorisation ability with preventive activities is based on the ability to memorise with some delay (as compared to time orientation or direct recall). One explanation for this is that the ability to memorise over a longer period can be a fundamental layer of further dimensions of cognitive functioning, as suggested by Serper et al. (2014).

Our findings have strong policy relevance, as the target group for the screening programme consists of people aged 60 to 70, and it is well known that a higher decrease in cognitive functioning can be expected during later life (Salthouse, 2009). Due to the demographic changes which have successively increased the share of elderly people in many developed countries, specific characteristics of this older population should be considered when implementing public (health) programmes. Although there has been an ongoing discussion about the benefits, costs and the (future) financial burden of preventive health programmes for society since Fries (1980), specific characteristics of the target group should obviously be accounted for from a normative point of view, given the goal of a specific policy. Hence, a receiver adequate way to provide information is needed, maybe based on the integration of different healthcare providers.

Although we provide evidence that a decrease in mental ability is a general phenomenon in the context of bowel cancer screening for the elderly and not driven by specific diseases, we cannot fully identify the point in the data generating process at which the decrease in memorisation ability determines screening behaviour. This would be a fruitful topic for future research. In addition, we cannot rule out the possibility that further unobserved factors that cannot be controlled for in our analysis are confounding our estimates. From both a theoretical and applied perspective, it

is also important to determine how to provide sufficient information to a policy programme's target group to simultaneously mitigate potential (cognitive) constraints among decision makers while still allowing individuals to make individual choices.

Appendix Chapter 3

3.A1 Impact of characteristics on decline in memorisation ability (Logit)

Variables	Effect	z-stat
memory	0.085***	(15.64)
educ_mid	-0.057	(-1.52)
educ_high	-0.119**	(-2.55)
wealth	-0.000***	(-2.73)
adl	0.047	(1.16)
health_2	-0.038	(-0.56)
health_3	-0.038	(-0.55)
health_4	-0.061	(-0.83)
ill1	-0.022	(-0.79)
ill2	-0.040	(-1.06)
ill3	0.020	(0.68)
ill4	-0.059	(-1.13)
ill5	-0.010	(-0.27)
ill6	0.007	(0.17)
female	-0.112***	(-3.77)
black	0.176	(1.51)
couple	0.010	(0.28)
smoker	0.089**	(2.03)
smoke_past	0.054*	(1.79)
organisation	-0.017	(-0.58)
phys_act2	-0.075	(-0.78)
phys_act3	-0.033	(-0.35)
phys_act4	-0.075	(-0.77)
life_exp	-0.000	(-0.58)
N	1564	
$pseudoR^2$	0.152	

Notes: Average marginal effects (including age fixed effects), robust standard errors, z-statistics (in parentheses)
* p<0.10, ** p<0.05, *** p<0.01.

Chapter 4

Heterogeneous Parameters and Detecting Selection Based on 'Unused Characteristics' in Private Health Insurance Markets[1]

4.1 Introduction

In addition to access to the National Health Service (NHS), around 11 % of the UK population have either private health cover for specific conditions (such as cancer), or broader coverage which includes complementary therapies and diagnostic tests (Cylus et al., 2015). There are several possible explanations for people's different private insurance requirements, based on the supply and demand sides of the insurance market. One common explanation is that risk preferences, or an individual's degree of risk aversion, influence the probability that an individual will buy insurance. Another explanation asserts that insurance companies do not take into account risk heterogeneity when calculating premiums, but customers take this into account when deciding whether to buy insurance. However, the converse may also be true: insurance companies use some risk-related customer characteristics to lower their exposure to high risk policyholders. During the last decade, a great deal of literature has emerged which attempts to identify the specific origins of selection and information asymmetries (IAs) in insurance markets (e.g., Finkelstein and McGarry, 2006; Fang et al., 2008). However, evaluating these issues is still an important

[1]This study is based on joint work with Martin Karlsson and Ben Rickayzen. See Karlsson et al. (2012) for a very early draft of our research.

topic in economic research, and little research has been done in this area before now, particularly in relation to UK data.

And individual's level of health directly affects their expected contribution to the economy, their happiness and their ability to participate within society at large. Providing healthcare for an entire population is becoming increasingly expensive, in part due to technical advances in treatments but also due to demographic change (i.e., an ageing population). In the face of restricted public budgets, private health insurance that partly or fully takes on coverage is a potential remedy (e.g., Arentz et al., 2012; Leidl, 2008). Since the seminal works by Akerlof (1970) and Rothschild and Stiglitz (1976) on market efficiency in insurance markets, there has been a preponderance of theoretical and empirical research into selection based on IAs. However, how to accurately and adequately measure IAs and their consequences remains unexplored. This is a very important area of research, since it directly affects the funding of a country's health and welfare system. A specific branch of the literature (Finkelstein and McGarry, 2006; Fang et al., 2008) concerns the indirect detection of IAs by harnessing 'unused characteristics', i.e., variables that insurance companies do not use to calculate risk premiums, but which are still available as observations for empirical analysis.

The aims of this chapter are threefold. The main contribution of this study is a technical one, and we first formally demonstrate that standard 'unused characteristics' approaches which allow the detection of IAs based on specific characteristics can lead to erroneous conclusions. This is the case if the parameter of a potential source of IA in a framework with two equations differs from an individual's risk/insurance status, i.e., whether the estimated coefficient of interest is driven by different parts of the population. Second, we create artificial data to emphasise this issue under different assumptions about the underlying data generating process that causes a selection effect in the insurance market. Third, taking this phenomenon into account, we provide empirical evidence of selection in the market for private health insurance using data for the English population over age 50.

We begin by describing the institutional background to the current healthcare system in England. We then provide a literature overview, with an emphasis on several commonly used tests to identify IAs. Following this, we show formally and by simulation that, under specific circumstances, tests based on two equations that try to detect IAs with 'unused characteristics' can be misleading. The empirical section of this chapter provides evidence by applying such tests, and also allowing for individual parameter heterogeneity by using a multilevel model. Our empirical analysis is based

on the English Longitudinal Survey of Ageing (ELSA), an individual-level dataset representative of the English population over age 50. ELSA has both a cross-sectional and longitudinal dimension. After discussing our findings, we draw conclusions and make suggestions for future research in this field.

4.2 Theoretical Considerations

4.2.1 Institutional Background

The population of England is entitled to free healthcare, provided by the National Health Service (NHS) through primary care (general practice) and secondary care (hospital-based care given through both NHS and Foundation Trusts). The guiding principle of the NHS is to make health services available to every citizen in need.[2] However, in practice, there are a number of treatments which are not available within the system. Most of these are excluded because they are viewed as being non-essential, but some are excluded for financial reasons.[3] In addition to the public provision of healthcare via the NHS, individuals can choose to top up their provision by purchasing private health insurance (PHI). This can be done on an individual basis or as part of a benefits package offered by employers.(Boyle, 2011) Private insurance covers services which duplicate those provided under the NHS (Kiil, 2012), but also provides cover for enhanced services such as faster access and wider choice. Insurers can freely determine the services they offer, but most packages cover surgery as an inpatient or day case, hospital accommodation, nursing care and inpatient tests. Since there is no regulation on products or pricing (Boyle, 2011), we can assume that the market for (voluntary) PHI in England is competitive. Although a competitive market should result in an actuarially fair risk premium due to the possibility of consumers switching between different contracts, we cannot draw any conclusions about the efficiency of this market, i.e., selection effects due to IAs which are not accounted for in the risk premium.

4.2.2 Detecting Information Asymmetries

Empirical evidence for the existence of IAs is mixed. Cohen and Siegelman (2010) provide a metastudy on testing for adverse selection in a wide range of insurance markets. They focus on the positive correlation approach and find a correlation

[2] http://www.nhshistory.net/a_guide_to_the_nhs.htm
[3] http://www.londonhealth.co.uk/nhs/index.html

between risk and insurance in some studies but not others. For example, evidence in health insurance markets appears strongly heterogeneous. Looking at studies which focus on the US market, they find evidence for both the existence of IAs and market efficiency. They also assert that it is necessary to distinguish between different kinds of IA. They conclude that it might be useful to evaluate the circumstances under which adverse selection does or does not arise. This perspective is particularly relevant from a policy perspective, given that we would like to be able to predict efficiency changes based on initial market conditions in the face of any institutional changes. In their work, Cohen and Siegelman (2010) mainly focus on an approach to detecting IA, developed by Chiappori and Salanié (1997), which is usually called the 'positive correlation test'. This test is still widely used today, despite ongoing developments in the field. The main thrust of the test is to jointly estimate two separate equations. The first captures the probability of buying an insurance contract given the information about an individual which an insurance company will use to calculate the risk premium. The second measures the correlation between these variables and the probability of the insurer making a loss on the contract. The error term in both equations covers all the information about both events which is not used for pricing purposes. If risk and insurance coverage are correlated, this is usually interpreted as indicating that a self-selection process is occurring based on unused variables. Hence, it is useful to estimate the correlation between both equations' error terms. This approach is often called the 'positive correlation test'. Formally, this approach can be described by the following equations, where I is an indicator of insurance status and R is an indicator of being at risk, while X is a matrix containing the variables used by the insurance company to calculate the risk premium:

$$I = X\phi + \epsilon \tag{4.1}$$

and

$$R = X\psi + \eta \tag{4.2}$$

Subsequent literature assesses the problem of multiple dimensions of private information in the context of detecting IA, which is also a focus of this chapter. As Finkelstein and McGarry (2006) (FMG) argue, the correlation between error terms in the 'Chiappori approach' is neither a necessary nor a sufficient condition for the

existence of IAs. The authors suggest that misleading results may arise if several characteristics have an impact on both dependent variables (some negatively, some positively) and effects cancel out on average. For example, in addition to an individual's class of risk, heterogeneity in consumer risk preferences might offset the correlation between the two equations' error terms. The authors assert that if an econometrician can identify such relevant information, and this information is not used by the insurer for pricing, then including this variable as an additional explanatory variable into equations (4.1) and (4.2) will make it possible to detect and separate out this kind of self-selection, despite the second relevant variable having an offsetting effect. This approach, which we call the 'unused characteristics' approach, is based on the following equations:

$$I = X\phi + Z\delta + \epsilon \tag{4.3}$$

and

$$R = X\psi + Z\beta + \eta. \tag{4.4}$$

where Z represents a matrix containing additional information about the insured but which is not used for pricing. The condition for recognising IAs is that any new variable included in the model affects the probability of both getting insurance and 'being at risk'. In their study, Finkelstein and McGarry use information that is assumed to be unknown to the insurer in the market for long term care (LTC) in the US.

The unused characteristics approach described above can also be useful if we are interested in selection in terms of pricing, without using 'unobserved' information. In contrast to the case mentioned above, Finkelstein and Poterba (2014) focus on a scenario where insurance companies observe, or could observe, relevant customer characteristics, but do not use this information when calculating their risk premium. Analysing the UK annuity market, they show that annuity purchases and the annuitant's mortality are regionally correlated. Assuming that regional information is not used to calculate the risk premium, this is interpreted as an indicator of adverse selection. Finkelstein and Poterba's results raise the question of why insurance companies do not tend to use this kind of information.

Cutler et al. (2008) look for selection within several insurance markets in the US, based on data from the Health and Retirement Study. They also use a dual-

equation model with insurance status and risk occurrence as dependent variables. While conditioning on variables used for insurance pricing, they also include some behavioural variables which are used to measure heterogeneity in risk preference (e.g., seatbelt usage, preventative activities) and individual risk behaviour (e.g., alcohol consumption), which are probably not available to an insurance company. Their findings suggest advantageous selection in the market for life and LTC insurance, but adverse selection for annuities.

Fang et al. (2008) (FKS) develop a similar approach that tries to reveal unused characteristics that drive selection in insurance markets. Assuming unused information is already partialled out, their approach is based on the regression model:

$$I = \alpha_1 + \alpha_2 R + \nu_i \tag{4.5}$$

followed by a regression which includes an unobserved variable, z:

$$I = \gamma_1 + \gamma_2 R + \gamma_3 z + \mu_i. \tag{4.6}$$

It can be shown that the regression based on (4.5) will result after applying the expectation operator into $\mathbb{E}(\hat{\alpha}_2) = \gamma_2 + \gamma_3\theta_{32}$, where θ is based on the auxiliary regression $R = \theta + \theta_{32}z + \lambda_i$. The detection of IAs is based on the difference between the estimates for α_2 and γ_2. Hence, the detected IAs induced by z are defined as $\mathbb{E}(\hat{\alpha}_2) - \mathbb{E}(\hat{\gamma}_2) = \mathbb{E}(\gamma_2) + \mathbb{E}(\gamma_3\theta_{32}) - \mathbb{E}(\gamma_2) = \mathbb{E}(\gamma_3\theta_{32})$. For advantageous selection, this difference will be negative, i.e., $(\gamma_2 > \alpha_2)$ if $\gamma_3 < 0$ and z is partially positively correlated with R. Unlike the approach suggested by FMG, the evidence is not directly based on comparing two different outcomes for z, but on a single coefficient and the partial correlation between z and R.

Both approaches are applied in the literature. For example, Cutler et al. (2008) apply the FMG approach and look for IAs within several insurance markets in the US, based on data from the Health and Retirement Study. Bolhaar et al. (2012) also implement the FMG framework to assess multidimensional asymmetric information in Ireland, whereas the approach developed by FKS is applied by Buchmueller et al. (2013), who analyse advantageous selection using Australian data.

There is little evidence on selection in the market for PHI in England. Propper et al. (2001) analyse the dynamics in the demand for PHI between 1978 and 1996 in the UK, using the Family Expenditure Survey. Controlling for consumer

characteristics and health service quality measures, they find that the availability of private healthcare facilities and cohort effects, which might indicate changes in tastes/attitude to PHI, are important factors in deciding whether to purchase PHI.

Wallis (2004) also looks for the determinants of demand for PHI in the UK, based on data from the British Household Panel Survey (BHPS). The author evaluates individuals' switching behaviours and focuses on characteristics which influence the probability of purchasing insurance and those which influence the individual cost of PHI (i.e., the risk premium). The study differentiates between consumer demand-side characteristics and supply-side factors that can influence insurance status, e.g., quality of service.

Another study using BHPS data was carried out by Olivella and Vera-Hernández (2013), who focused on adverse selection in the market for PHI, using hospitalisation as a measure for being at risk. Assuming that an individual's health status is independent of receiving PHI as a fringe employment benefit, their results suggest the existence of adverse selection in the PHI market in England.

Until now, there has been no empirical investigation into whether specific sources of selection exist in the English PHI market.

4.2.3 Using Unused Variables

Finkelstein and McGarry (2006) argue that the standard positive correlation test suggested by Chiappori and Salanié (1997) will fail if there is more than one characteristic that is not used for pricing purposes but which affects insurance status and risk of loss to the insurance company. These characteristics can offset the correlation between the equations' error terms.

The advantage of an unused characteristics approach is that it allows the identification of specific characteristics which can, from a theoretical perspective, be assumed to be a source of IAs, even if the positive correlation test does not reveal as much. It is perfectly reasonable to interpret the coefficients in these models if we have an underlying theory as to why these characteristics should be correlated with both insurance and health status. Thus the suggested approach is very helpful when we want to assess whether a specific characteristic is a source of IA in a setting with multidimensional private information. However, we do not know how this approach performs if we allow for another source of heterogeneity. More precisely, it is not clear how evidence based on this approach changes if we allow the outcomes of an unused characteristic to be heterogeneous across the risk pool. This is a very important issue, and parameter heterogeneity has not been discussed so far in this

context.

The motivation behind this question stems from the fact that conclusions about selection effects are derived based on the relevance of the variables included in both equations. This approach can, however, be problematic if we find an additional variable to be relevant in our framework and we do not have a good a priori theory about the mechanism which relates this variable to risk and insurance status. This is the case since a variable might, on average, control for relevant factors which explain, for example, insurance probability and risk situation, but it is possible that these estimated coefficients are driven by different parts of the sample. We suggest that if individuals with a shared characteristic are heterogeneous in outcomes of this characteristic (i.e., marginal changes in risk and insurance probability due to an unused variable being negatively or positively related) an unused characteristics approach can lead to erroneous conclusions.

From this, it follows that an individual's expected risk from an insurance perspective, conditioned on certain characteristics (which are not used to calculate the risk premium), can equal the population's expectation of risk, even though such characteristics are related to both risk and insurance probability, which would usually be interpreted as an indicator of IAs. From a policy perspective, this is a crucial issue for both judging the efficiency of an insurance market and predicting changes in welfare when planning to implement new policies (e.g., regulation of contracting health services). To describe this issue formally, we simplify equations (4.3) and (4.4) without loss of generality, assuming that there is no used 'observable' information, X, included in either equation, allowing for just one 'unused' variable, z:

$$\mathbb{E}(I|z) = \mathbb{E}(\delta z + \epsilon)$$

and

$$\mathbb{E}(R|z) = \mathbb{E}(\beta z + \eta)$$

where estimates $\hat{\beta} > 0$ and $\hat{\delta} > 0$ would usually be interpreted as an indicator of adverse selection. We now relax the assumption of parameter homogeneity and assume that every parameter is defined for every individual, i, with the population means $\mathbb{E}(\beta_i) = \mu_\beta > 0$ and $\mathbb{E}(\delta_i) = \mu_\delta > 0$. We further assume, for simplification purposes, that $\text{Cov}(z_i, \beta_i) = 0$, $\text{Cov}(z_i, \delta_i) = 0$ and $\text{Cov}(\eta_i, \epsilon_i) = 0$.

Based on this model, we now evaluate the risk position of the sub-population being insured ($\delta_i z_i + \epsilon_i > 0$):

$$\mathbb{E}(R_i|I_i > 0) = \mathbb{E}(\beta_i z_i|\delta_i z_i > -\epsilon_i)$$

Given the population means μ_β and $\mathbb{E}(z_i) = \mu_z$, this equation can be rewritten as

$$\begin{aligned}\mathbb{E}(\beta_i z_i | \delta_i z_i > -\epsilon_i) =& \mu_\beta \mu_z \\ &+ \mu_\beta \times \mathbb{E}(z_i - \mu_z | \delta_i z_i > -\epsilon_i) \\ &+ \mu_z \times \mathbb{E}(\beta_i - \mu_\beta | \delta_i z_i > -\epsilon_i) \\ &+ \mathbb{E}((\beta_i - \mu_\beta)(z_i - \mu_z) | \delta_i z_i > -\epsilon_i).\end{aligned} \tag{4.7}$$

We want to find the circumstances under which $\mathbb{E}(R_i) \geq \mathbb{E}(R_i | I_i > 0)$, since this is when the unused characteristics approach has falsely detected adverse selection. The first and second terms in decomposition (4.7) are positive based on our assumptions. Hence, $\mathbb{E}(\beta_i z_i | \delta_i z_i > -\epsilon_i) \leq \mathbb{E}(R)$ can be true if $\mathbb{E}(\beta_i - \mu_\beta | \delta_i z_i > -\epsilon_i) < 0$ or $\mathbb{E}((\beta_i - \mu_\beta)(z_i - \mu_z) | \delta_i z_i > -\epsilon_i) < 0$. Therefore, we would need $\mathrm{Cov}(\beta_i, \delta_i) < 0$, and thus $\mathbb{E}(\beta_i | \delta_i z_i > -\epsilon_i) < \mu_\beta$ to offset the other terms in the decomposition.

This condition requires the expectation that, for an individual whose β_i is smaller than the population mean, μ_β, the individual coefficient, δ_i, will be higher than the estimated population mean, μ_δ, and vice versa.

Clearly our framework requires $\mathrm{Cov}(\beta_i, \delta_i) > 0$ in the case where $\mu_\beta < 0$ and $\mu_\delta > 0$ (or vice versa), i.e., wrongly detected advantageous selection using the unused characteristics approach. If the correlation between β_i and δ_i is strong enough, the suggested direction of selection may even be the opposite of what it should be.

To solve this problem, it is possible to compare insurance and risk status at the individual level, as done by Chiappori and Salanié (1997), by calculating the correlation between the error terms. As Finkelstein and McGarry (2006) argue, this approach is not helpful in all contexts, since some correlations can cancel out.

Does the problem of parameter heterogeneity also affects the approach suggested by Fang et al. (2008)? Remember, this model emphasizes the partial regression coefficient $\gamma_{Iz|R} \neq 0$ and the difference between $\gamma_{IR|z}$ and the correlation α_{IR} when detecting IAs. This is very similar to the FMG approach, since both approaches assume that z is a determinant of both risk and insurance status. However, the selection in both the FMG and FKS approaches is assumed to be based on constant coefficients, which by definition do not reflect parameter heterogeneity. Hence, the above discussion of parameter heterogeneity is also relevant in this context. The following section provides an empirical comparison between both approaches in terms of the effect of correlated coefficients.

4.2.4 Econometric Model

The need to determine parameter heterogeneity at the individual level (δ_i, β_i) imposes stronger requirements on the data than analysis where both the FKS and FMG approaches are usually applied. To be able to estimate coefficients for each individual in the data, we need to add degrees of freedom at the unit level where heterogeneity occurs. Therefore, we rely on a panel data setting in the following. We are aware of one other study by (Bolhaar et al., 2012) that identifies unused characteristics in the context of PHI using panel data. However, while they use dynamic panel data estimators to distinguish between short- and long-term determinants of selection, we emphasise the role of different outcomes from the unused characteristics, z.

To derive the following equations, we now assume a longitudinal data structure with $t = 1, .., T$ periods, where δ and β represent vectors containing our $i = 1, .., N$ coefficients, δ_i and β_i.

$$I_{it} = \delta^{'} z_{it} + X\phi + \epsilon_{it} \tag{4.8}$$

and

$$R_{it} = \beta^{'} z_{it} + X\psi + \nu_{it} \tag{4.9}$$

Our analysis is based on estimating the impact of z_{it} on R_{it} and I_{it} while jointly controlling for other observable characteristics which are used to calculate the risk premium, captured by the matrix X. The coefficients of z_{it} are allowed to be individual-specific, but we impose the assumptions of parameter constancy on all other variables in the model, as suggested by standard unused characteristics approaches. To see whether conclusions based on these approaches can be misleading due to parameter heterogeneity, we will calculate $\text{Corr}(\hat{\beta}_\text{i}, \hat{\delta}_\text{i})$.

When assessing the role of heterogeneity in selection into the English health insurance market we use a linear probability model, which is a simplification of the framework provided above, since our dependent variables are no longer continuously defined.

In our comparative analysis, we apply both unused characteristics approaches by pooling the data to compare their results. In an additional specification, we allow for individual parameter heterogeneity and estimate equations (4.8) and (4.9) using multilevel models, where β_i and δ_i are supposed to be determined at the individual level. We provide an example with a variable that is supposed to induce selection into the insurance market and test whether the estimates for β_i and δ_i reveal a corre-

lation structure that suggests a bias in the interpretation of the standard approaches, i.e., both $\mathrm{E}(\beta_\mathrm{i}) \times \mathrm{E}(\delta_\mathrm{i}) > 0$ based on pooled regressions but $\mathrm{Corr}(\hat{\beta}_\mathrm{i}, \hat{\delta}_\mathrm{i}) < 0$ when accounting for parameter heterogeneity, or $\mathrm{E}(\beta_i) \times E(\delta_i) < 0$ but $\mathrm{Corr}(\hat{\beta}_\mathrm{i}, \hat{\delta}_\mathrm{i}) > 0$. Such a finding, ignoring the correlation structure between δ_i and β_i, is usually interpreted as indicating a selection mechanism due to IA. When applying the FKS approach, we derive the direction and degree of selection based on the difference between the partial correlation between I and R before and after an unused characteristic is included in the specification. We then run regressions allowing for parameter heterogeneity by regressing I and R on z using multilevel models, fixing the parameters δ and β at the individual level. All other variables, X, included in the specifications are assumed to have a fixed impact on the outcome variables, as usually assumed by unused characteristics approaches. We use the correlation between $\hat{\delta}_i$ and $\hat{\beta}_i$ as a measure to test whether an offset is occurring in the standard approaches, as described above.

Before analysing the phenomenon of parameter heterogeneity in the context of selection into English PHI market, we first assess how the approaches discussed so far perform under different assumptions via simulation.

4.3 Empirical Implementation

4.3.1 Simulations

In order to compare the outcomes for both unused characteristics approaches under parameter heterogeneity at the individual level, we create artificial data by imposing different assumptions about variables, parameter heterogeneity and the correlation between variables. Panel data with 1,000 cross-sectional observations and five time-series units is generated with assumptions about $\mathbb{E}(\delta_i)$, $\mathbb{E}(\beta_i)$ and $\mathrm{Corr}(\hat{\beta}_\mathrm{i}, \hat{\delta}_\mathrm{i})$. Assuming that other information, X, used to calculate the risk premium is already partialled out before estimation, the simulations are based on the following structural equations:

$$I_{it} = \delta_i z_{it} + \epsilon_{it} \tag{4.10}$$

and

$$R_{it} = \beta_i z_{it} + \eta_{it} \tag{4.11}$$

where z_{it}, ϵ_{it} and η_{it} are standard normal random variables, whereas δ_i and β_i are random variables that vary between each of the 100 simulations. The population

means, standard deviations and degree of correlation between the parameters δ_i and β_i vary between different scenarios. We use the generated data to run 6 times 100 regressions for each scenario and calculate the correlation between the estimated random coefficients.

We divide our analysis into four separate blocks, which are further divided into different sub-scenarios. In our simulations, we assume that IAs are solely driven by an unused characteristic, z. We also assume that observable characteristics which are allowed to be used to calculate the risk premium are already partialled out.

Column 1 (see tables in Appendix 4.A1 to 4.A4 of this chapter) shows the degree of selection within the insurance market. Since, by definition, selection is solely driven by z, we choose the regression coefficients of I on R ($\hat{\alpha_2}$) as a measure to detect IAs. Columns 2 to 4 show further results of the FKS approach, estimating specifications based on equations (4.5) and (4.6). Columns 5 and 6 show evidence based on regressions reflecting the results of FMG approach, while columns 7 to 9 show results for the multilevel model with random coefficients for each cross-sectional unit. We now compare the extent to which both approaches reveal IAs associated with z.

Scenarios 1a to 1c in Appendix 4.A1 represent the case where both the FMG and FKS approaches suggest no selection due to z, i.e., $E(\delta_i) = E(\beta_i) = 0$, but a positive selection due to $\text{Corr}(\hat{\beta}_\text{i}, \hat{\delta}_\text{i}) < 0$ (scenario 1a) and a negative selection in scenario 1c. Although both types of selection are well detected by a regression of I on R (column 1), the association with z is not detected using either standard unused characteristics approach. We see that the outcomes in columns 3 to 6 are unaffected by the joint variation of β_i and δ_i.

To investigate the role of the standard deviation of β_i and δ_i (Appendix 4.A2), we let it vary between 0.5 (row 2a) and 1.5 (row 2c). Obviously, the higher the standard deviation of the coefficients, the greater the role of $\text{Corr}(\hat{\beta}_\text{i}, \hat{\delta}_\text{i})$ in the selection mechanism. This implies that, given the levels of β_i and δ_i for z, a higher coefficient variation positively affects selection. Again, neither the FKS nor the FMG approach detect this kind of selection, since columns 3, 5 and 6 remain stable.

In alternative scenarios (scenarios 3a-3c in Appendix 4.A3), we try to determine what happens if there is a positive correlation between z and R but a negative correlation between z and I. Do these associations necessarily induce advantageous selection? Increasing the association $\text{Cov}(\hat{\beta}_\text{i}, \hat{\delta}_\text{i})$ from 0 (row 3a) to 0.6 (row 3c) again reveals that the point estimates of the standard unused characteristics approaches are very similar across the scenarios, but the issue of parameter heterogeneity is

completely ignored, i.e., both approaches suggest advantageous selection. This is also revealed for the FKS approach in column 3, where the difference $\mathbb{E}(\hat{\alpha}_2) - \mathbb{E}(\hat{\gamma}_2)$ is found to be negative across all scenarios. Columns 5 and 6 have opposite signs, which is interpreted as an advantageous selection in the FMG approach. In contrast, negative selection in the insurance market is revealed by the increasing positive correlation between the estimated random coefficients (row 3b and 3c in column 9).

Scenarios 4a-4c (Appendix 4.A4) show the implications if there is a positive correlation between both z and R and z and I, i.e., the unused characteristics approach suggests adverse selection, but $\text{Corr}(\hat{\beta}_\text{i}, \hat{\delta}_\text{i}) < 0$ suggests offsetting. Adverse selection is identified under both the FKS ($\mathbb{E}(\hat{\alpha}_2) - \mathbb{E}(\hat{\gamma}_2) > 0$) and FMG approaches (positive estimates in columns 5 and 6), despite decreasing $\text{Cov}(\hat{\beta}_\text{i}, \hat{\delta}_\text{i})$ from 0 (row 4a) to -0.6 (row 4c), which actually implies advantageous selection in scenarios 4b and 4c. This advantageous selection is suggested by the negative correlation between our random coefficients (rows 4b and 4c in column 9).

4.3.2 ELSA Data

Our empirical exploration of this idea is based on ELSA, which is a representative individual-level dataset for England's 50+ population. The ELSA dataset contains a broad range of information on each individual's health and financial circumstances, together with overall demographics, which makes it an ideal source to model both economic decisions and health-related characteristics.

For our analysis we use both the cross-sectional and longitudinal dimensions of the ELSA survey to ensure the period we are using captures the time from 2002 to 2013. For our analysis, we restrict the data to individuals aged 90 or younger, since we cannot verify the actual age of people over age 90. Due to the longitudinal nature of our dataset, we only analyse individuals who were eligible for an interview during all six ELSA waves. In our empirical analysis, we apply the sample weights provided with the ELSA data that account for attrition and which make the sample we use representative for the first wave.

We use self-assessed health as the main dependent variable as a measure for being at risk. Although this information can be subjective, we assume it to be a reasonable indicator, since it does not only capture observable information (which we control for in our analysis), but all information that can affect future demand for healthcare which cannot be accounted for by using only observable and objective health data. Based on their findings, Idler and Benyamini (1997) argue that a global health rating "... represents an irreplaceable dimension of health status and in fact

that an individual's health cannot be assessed without it." Other studies into IAs, such as those by Doiron et al. (2008) and Bolin et al. (2010), also use self-assessed health to capture individual risk. Hence, we assume that self-assessed health is a suitable measure for our purposes.

ELSA provides the commonly used 5-point self assessed health (SAH) measure, which we collapse into binary variables that we call 'high risk' (HR) and 'low risk' (LR) (see descriptive statistics in table 4.1[4]). HR captures having 'fair' or 'poor' health, while LR captures 'excellent', 'very good' and 'good' health. HR is used in our analysis to capture information when an individual poses a relatively high health risk from an insurance company's perspective. The second main dependent variable is a dummy variable, PHI, which equals 1 if an individual has private health insurance and 0 otherwise. We exclude people who only have PHI cover as part of an employee benefits package offered by their employer. This is because the way in which such group cover is purchased by employers and the way in which it is priced, are both very different from the approach adopted for individual policies.

Table 4.1 Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
privins	0.126	0.332	0	1	31431
HR	0.266	0.442	0	1	31431
time_avail	0.774	0.418	0	1	31431
female	0.565	0.496	0	1	31431
age	67.629	8.958	50	89	31431
nwhite	0.015	0.121	0	1	31431
couple	0.690	0.463	0	1	31431
children	0.883	0.322	0	1	31431
educ	0.688	0.463	0	1	31431
working	0.289	0.453	0	1	31431
not_work	0.112	0.316	0	1	31431
retired	0.599	0.49	0	1	31431
smoke_now	0.129	0.335	0	1	31431
smoke_past	0.5	0.5	0	1	31431
ill1	0.431	0.495	0	1	30314
ill2	0.181	0.385	0	1	30314
ill3	0.451	0.498	0	1	30314
ill4	0.092	0.29	0	1	30314
ill5	0.116	0.32	0	1	30314
ill6	0.289	0.453	0	1	30314

[4]Information about variable definitions can be found in Appendix 4.A5

As previously mentioned, it is important for our analysis to assume that the econometric model contains all the relevant information used by an insurance company to calculate the risk premium. Since the people in our sample have PHI with different suppliers, we cannot provide a general framework for calculating these premiums. We therefore have to make reasonable assumptions and try to stay as close to the existing literature as possible. We assume that the following variables are used by insurance companies: age, sex, smoking (history), employment status (i.e., in the labour market, not in the labour market, retired), education (where 'no qualification' is the reference group), race, family status (indicator for being married or cohabiting) and whether an individual has children. We also use dummy variables for the government office region the respondent is living in. Since we have detailed information on each individual's health, we also use indicator variables capturing a broad range of self-reported illness categories which are available in the data and are assumed to be used to calculate risk premiums. This is important, since insurers require applicants to provide detailed information about their past and present health status (cf. Boyle, 2011).

For the unused characteristics, we take a variable from the literature which is known to drive selection within the PHI market yet which cannot directly be assessed by an insurance company and is therefore not available to calculate premiums. From a health economics perspective, the restrictions on an individual's time are of great interest, since the decision to take out PHI in the UK is known to depend on waiting times for healthcare (e.g., King and Mossialos, 2005; Johar et al., 2013). As the waiting times in the English healthcare market are usually much shorter for those with PHI than under the NHS, an individual's available time and the corresponding opportunity costs will determine whether they take out PHI or not. If a patient is willing and able to wait longer for treatment, there is less incentive to opt into the PHI market.

Hence, we directly utilise self-assessed information on respondents' available time. A respondent's subjective level of available time may reflect low opportunity costs when facing a relatively long waiting time for healthcare services. As shown above, this directly decreases demand for PHI. Therefore, we would expect people with a relatively large amount of available time to take out less insurance. We also expect that the relationship with an individual's health risk status will be negative if a larger amount of available time affects health-related decisions that increase the decision maker's health stock, lowering their health risk. However, actual health-related decisions depend on other factors such as an individual's preferences or the

urgency of a specific case. Therefore, we regard the assessment of such a relationship as mainly an empirical question.

In order to capture a respondent's level of available time, we use their answer to the question 'Do you have enough time to do everything?' to see whether this factor drives selection in the PHI market and whether there is individual heterogeneity in its outcomes. The ELSA survey offers six different responses to this question, ranging from 'strongly agree' and 'strongly disagree'. To give our resulting estimates a proper meaning, we collapse 'agree' and 'do not agree' into one binary variable that takes a value of 1 if the respondent's answer to the question reflects agreement and 0 otherwise. We provide results for three different specifications. In the baseline specification, we condition on variables that are assumed to be used to calculate the risk premiums for PHI contracts. We then further use time dummies to rule out changes over time that may affect our estimates. Finally, we also include variables capturing diseases from certain health domains.

4.3.3 Results and Discussion

The coefficients from the linear probability model[5] based on the FMG approach (see specification 1 from Table 4.2) show that an individual's available time is negatively correlated with that individual's health risk status (-0.014) and their ownership of PHI (-0.076)[6]. Taken together, this would usually be interpreted as adverse selection (due to z) under FMG. However, as can be seen in column 3 under the FKS approach, there is a negative correlation between health risk and insurance overall. This empirical finding is also predicted by Olivella and Vera-Hernández (2013), who discuss the issue of IAs in the UK's PHI market. However, we are not interested in the overall degree of IAs, instead focusing on a specific characteristic, z, that may imply selection.

The FKS approach also indicates $\text{Cov}(R, z) < 0$ and $\text{Cov}(I, z) < 0$, although the impact of health risk on I is very similar once our unused variable z is conditioned on (column 4), and the difference is not statistically significant. However, the literature does not usually test whether the differences are statistically significant (e.g., Buchmueller et al., 2013; Finkelstein and Poterba, 2014). When focusing on the estimates

[5]Despite theoretical concerns over the interpretation of the coefficients in an LPM (e.g., Wooldridge, 2003), we find that the LPM fits our data well. Its coefficients are very similar to the marginal effects derived from a probit model, both in terms of economic relevance and statistical significance.

[6]The results for other control variables can be found in the Appendix 4.A6, but are not of interest for our analysis, since we assume they are used when calculating the individual's risk premium.

Table 4.2 Estimates ELSA data

Spec.	N	FMG		FKS			RC-Model		
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		$\hat{\beta}$	$\hat{\delta}$	$\hat{\alpha}_2$	$\mathrm{E}(\hat{\alpha}_2 - \hat{\gamma}_2)$	$\hat{\gamma}_3$	$\mathrm{E}(\hat{\beta}_i)$	$\mathrm{E}(\hat{\delta}_i)$	$Corr(\hat{\beta}_i, \hat{\delta}_i)$
1	31431	-0.014**	-0.076***	-0.034***	0.001	-0.017***	-0.011*	-0.067***	-0.061***
		(0.026)	(0.000)	(0.000)	0.829	(0.008)	(0.053)	(0.000)	(0.000)
2	31431	-0.014**	-0.075***	-0.033***	0.001	-0.017***	-0.013**	-0.063***	-0.060***
		(0.024)	(0.000)	(0.000)	0.830	(0.008)	(0.017)	(0.000)	(0.000)
3	30314	-0.014**	-0.046***	-0.037***	0.001	-0.016**	-0.014**	-0.040***	-0.058***
		(0.027)	(0.000)	(0.000)	0.893	(0.013)	(0.017)	(0.000)	(0.000)

Notes: The columns for each specification show results of unused characteristics approaches. Estimates in columns (1) to (5) are based on LPMs. Columns (1) and (2) are estimates from the FMG approach represented by the structural equations (4.3) and (4.4), whereas columns (3) to (5) reflect the estimates based on the structural equations, (4.5) and (4.6), suggested by FKS. Coefficients of other used characteristics are not shown above, but detailed regression results can be found in Appendix 4.A6. An F-test is used to test whether the coefficient of R (based on the specification in column 3) equals the coefficient of R after including z into the model. Columns (6) and (7) show the coefficients from a multilevel model where the explanatory variable 'available time' takes a random coefficient and the effect of all other variables is fixed. Finally, column (8) shows the correlation between the individual parameters $\mathrm{Corr}(\hat{\beta}_i, \hat{\delta}_i)$. Standard errors clustered at the individual level; t-statistics in parentheses. * p <0.10, ** p <0.05, *** p <0.01

in our multilevel model, we also find that the estimates for the fixed part of our model (columns 6 and 7) are very similar to the FMG approach. However, calculating the correlation between the estimates for our individual coefficients, $\hat{\delta}_i$ and $\hat{\beta}_i$, reveals that they are negatively correlated (-0.061). This means that although the impact of z on I (domain 1) and HR (domain 2) is negative, suggesting adverse selection, the impact on domain 1 is actually highest for the coefficients for which it is lowest in domain 2, and vice versa. As we use variation over time in our estimation strategy, we would like to see whether changes over time that are identical for the population as a whole are affecting our estimates. As can be seen in table 4.2, the estimates for specification 2 are nearly identical to those for specification 1. The same is true for specification 3 if the illness variables are included[7]. However, we see that the impact of z on PHI greatly decreases, which suggests that an individual's health is a strong predictor for uptake of PHI. This is consistent with the existing literature, since it is known that, in the context of the English NHS, healthier individuals tend to have a higher demand for PHI (Olivella and Vera-Hernández, 2013). Nevertheless, the main relationships revealed in specification 1 do not change. Although the unused characteristics approaches suggest adverse selection, this interpretation is clearly offset by a negative correlation with the coefficients of interest. Hence, we find robust evidence for the role of an individual's available time on health risk and demand for PHI in England. What is the explanation for our findings? As mentioned above, economic theory suggests a negative relationship between someone's available time and their desire to purchase PHI. However, there is a relatively small, but statistically significant, impact on health risk. Our findings can be explained as follows. Some individuals who, in relative terms, have a substantial amount of time are relatively healthy (property (a)) and take out more insurance (property (b)), while

[7]Note that the number of observations is slightly lower in this specification, since some survey respondents did not answer this question.

other individuals with the same characteristic are relatively unhealthy (property (c)) but take out less health insurance (property (d)). The negative correlation between available time, z, and both insurance and health status is found in the data, because the effects of properties (a) and (d) dominate the overall correlation between z and the outcomes. This heterogeneity at the individual level is captured with individual coefficients, and cannot be detected using the cross-sectional data normally used when applying unused characteristics approaches. Although this chapter assumes a specific correlation structure, $\text{Cov}(\delta_{\text{i}}, \beta_{\text{i}}) \neq 0$, for the outcomes of z for both R and I, one may also consider a correlation $\text{Cov}(\delta_{\text{a}}, \beta_{\text{a}}) \neq 0$ on an arbitrary level, a, that can confound the interpretation of a standard unused characteristics approach.

Our empirical findings from the ELSA data show that our idea is not simply a theoretical artefact, but can be found in the real world and should be carefully accounted for when detecting the data generating process of selection within insurance markets. Hence we conclude that, although a variable can be correlated with insurance and health status, the relationship does not necessarily tell us anything about the importance of selection in insurance markets or IAs. This is because the estimated coefficients of such a variable (indicating IA) can be the result of contributions from different parts of the population. In addition, even if no correlation is found between z and both I and R, there might still be a selection mechanism driven by z which can incorrectly be hidden when the standard unused characteristics approaches are applied. We believe our approach will be of great importance when empirically determining whether a specific characteristic is the source of adverse/advantageous selection within an insurance market without a clear theory about the underlying selection mechanism. In such cases it may well be that pure randomness in the outcomes of such a variable is falsely interpreted as selection in the market of interest, as demonstrated above.

Our findings show that, as two different sources of private information can offset the correlation between the errors in the approach by Chiappori and Salanié (Finkelstein and McGarry, 2006), including private information within an unused characteristics framework can lead to erroneous conclusions about the interpretation of the selection mechanism if said mechanism is heterogeneously associated with both risk and insurance status.

4.4 Conclusion

This chapter provided an overview of commonly used testing procedures for detecting selection in insurance markets, focusing on their strengths and weaknesses. We argued that, although the classical positive correlation test might lead to erroneous conclusions about selection, it still has the advantage that it uses cross-equation correlations of the residuals acquired at the individual level. We also showed that standard unused characteristics approaches (e.g., Finkelstein and McGarry, 2006; Fang et al., 2008), which are often used to identify specific sources of IAs, can be problematic if evidence is wrongly based on mean coefficients. We also provided empirical findings based on simulations and for the English PHI market.

To emphasise this potential problem, we formally discussed the circumstances that can lead to erroneous conclusions by allowing for individual heterogeneity in parameters. We demonstrated the relevance of this finding through simulations imposing different correlation structures between an unused characteristic, z, insurance status and risk, while also allowing for individual parameter heterogeneity in the data generating process. The results show that standard unused characteristics approaches do not identify this kind of heterogeneity, and thus a detected source of adverse selection may indeed be a source of advantageous selection if the individual coefficients are negatively correlated. The same phenomenon can obviously be found under certain correlation structures between the parameters if adverse selection or even no selection is detected.

Our empirical implementation used the English PHI market as an example, in combination with the unused characteristic 'available time', which is not directly assessable by insurance companies. This variable may reflect opportunity costs of waiting times in the healthcare sector, resulting in demand for PHI. Our findings show that individual parameter heterogeneity is also relevant to real markets. Although adverse selection within the insurance market can be empirically detected, this adverse selection should be interpreted carefully, since the estimated parameters may be strongly negatively correlated due to the heterogeneous outcomes for the variable of interest. Again, the estimated 'mean' coefficient of such variables can be driven by different parts of the population and do not allow a meaningful interpretation of the underlying selection mechanism. When interpreting the empirical findings, it must be remembered that our aim is to focus on one specific source of selection. The total degree of selection within the insurance market is beyond the scope of this study.

Since the relevance of parameter heterogeneity is an empirical question, and a

general conclusion for other markets and characteristics cannot be provided, anyone wanting to identify a specific source of selection within insurance markets should test this possibility. Our findings are important for analysing the efficiency of insurance markets. They will be of interest to both the insurance industry and policy makers, and should be accounted for when predicting outcomes of structural changes to insurance policies or the overall design of the insurance market. For example, the expected impact of new regulations on calculating risk premiums based on characteristics such as ethnicity or gender might be of interest in this context. If one wishes to assess counterfactual situations about policies to regulate insurance markets, and parameter heterogeneity of the variable of interest is an issue, then its impact on selection within the insurance market should obviously be accounted for to achieve a market that works as expected.

Our findings are particularly relevant in the case of unused variables for which, a priori, we cannot assume any specific relationship with either insurance or health status. In this case, it is essential to be very careful about potential parameter heterogeneity, because even random outcomes of such unused characteristics may be falsely interpreted as being sources of selection in insurance markets and market inefficiencies.

We do not make any claim about the implications for welfare in private insurance markets based on our findings. Nevertheless, the economic implications of our ideas should be taken into account in welfare analysis if parameter heterogeneity is an issue, since they directly affect the interpretation of which kind of selection is being identified in the observed market. We also leave the implications of our findings for optimal policy design open for further consideration.

In terms of empirical applications, we assume that a subjective health risk variable is a good indicator of individual health status, but we do not know whether it is also a good measure for future healthcare uptake. Hence, we make the implicit assumption in our analysis that people with a relatively low (or relatively high) self-assessed health status are correlated with a higher (or lower) probability of making a health insurance claim. Future research should evaluate whether the robustness of this assumption can be supported when using objective data about healthcare utilisation (e.g., number of doctor visits) or, even better, treatment costs.

Finally, note that we do not explain why an unused characteristic, z, is (or should be) expected to have a heterogeneous impact on both health risk and the decision to take out insurance. We simply make the claim that, if this is the case, then the commonly applied unused characteristics approaches can be as misleading as the

Chiappori approach (no correlation of the error terms) when two different characteristics offset each other (Finkelstein and McGarry, 2006). One interesting area for future research would be to allow for heterogeneity at another level, a, in the association between an unused characteristic, z, and both R and I. A general condition such as $\text{Cov}(\delta_{\text{a}}, \beta_{\text{a}}) \neq 0$ may also lead to erroneous conclusions based on unused characteristics approaches and would be assessable even using cross-sectional data.

Appendix Chapter 4

4.A1 Simulation 1

Scenario	Parameter	Selection	FKS			FMG		RC-Model		
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		$\hat{\alpha}_2$	$\hat{\gamma}_2$	$\mathrm{E}(\hat{\alpha}_2 - \hat{\gamma}_2)$	$\hat{\gamma}_3$	$\hat{\beta}$	$\hat{\delta}$	$\mathrm{E}(\hat{\beta}_i)$	$\mathrm{E}(\hat{\delta}_i)$	$Corr(\hat{\beta}_i, \hat{\delta}_i)$
1a	$\mathrm{E}(\delta_i)$=0, $\mathrm{E}(\beta_i)$=0, $\mathrm{Cov}(\beta_i, \delta_i)$=-0.3	-0.150 [0] [100] [0]	-0.150 [0] [88] [0]	0.000 [46] [0] [0] $[46]^{n-}$ $[54]^{n+}$	0.000 [48] [2] [4] $[0]^{-}$ $[0]^{+}$	0.003 [48] [2] [4] $[0]^{-}$ $[1]^{+}$;	-0.001 [49] [2] [4] $[43]^{n-}$ $[57]^{n+}$	0.001 [51] [3] [1]	-0.001 [48] [1] [3]	-0.245 [0] [100] [0]
1b	$\mathrm{E}(\delta_i)$=0, $\mathrm{E}(\beta_i)$=0, $\mathrm{Cov}(\beta_i, \delta_i)$=0	0.000 [48] [1] [3]	0.000 [47] [0] [0]	0.000 [48] [0] [0] $[48]^{n-}$ $[52]^{n+}$	0.000 [48] [4] [3] $[0]^{-}$ $[0]^{+}$	0.003 [48] [2] [4] $[0]^{-}$ $[1]^{+}$;	0.000 [48] [4] [3] $[48]^{n-}$ $[52]^{n+}$	0.001 [51] [3] [1]	-0.001 [47] [4] [3]	-0.001 [47] [2] [2]
1c	$\mathrm{E}(\delta_i)$=0, $\mathrm{E}(\beta_i)$=0, $\mathrm{Cov}(\beta_i, \delta_i)$=0.4	0.200 [100] [0] [100]	0.199 [100] [0] [88]	0.000 [54] [0] [0] $[54]^{n-}$ $[46]^{n+}$	0.000 [50] [3] [3] $[0]^{-}$ $[0]^{+}$	0.003 [48] [2] [4] $[2]^{-}$ $[0]^{+}$;	0.000 [51] [4] [5] $[61]^{n-}$ $[39]^{n+}$	0.001 [51] [3] [1]	-0.001 [45] [4] [4]	0.324 [100] [0] [100]
Simulations		100	100	100	100	100	100	100	100	100
$N \times T$		1000 × 5	1000 × 5	1000 × 5	1000 × 5	1000 × 5	1000 × 5	1000 × 5	1000 × 5	1000 × 5

Notes: Each scenario reflects the combination of $\mathrm{E}(\delta_i)$ and $\mathrm{E}(\beta_i)$ with the correlation structure $\mathrm{Cov}(\beta_i, \delta_i)$, σ_β and σ_δ set to 1. Column (1) indicates whether there is a significant correlation between insurance (I) and risk (R), indicating IA. Columns (2) to (4) show estimates of the structural equations, (4.5) and (4.6), of the FKS approach. Column (3) can be interpreted as the degree of IA associated with z (the 'omitted variable bias') suggested by the FKS model: An F-test is used to test whether the coefficient of R (based on the specification in column (2)) equals the coefficient of R in column (1). Columns (5) and (6) are based on estimates of the equations (4.3) and (4.4) from the FMG approach. Column (5) provides the estimate of the effect of z on R in the FMG approach, and column (6) shows its effect on I. These estimates are based on least squares estimates. Columns (7) and (8) show coefficients from a multilevel model where parameters δ and β are estimated for each individual. Standard errors are clustered at the individual level. Column (9) shows the correlation between $\hat{\beta}_i$ and $\hat{\delta}_i$. [] indicate counts from 1 to 100. [] on the right side of coefficients counts the number of positive values. [] on the left hand side under a coefficient counts the number of cases a coefficient is both negative and significantly different from zero (5 % level); [] on the right side under a coefficient suggests the number of positive coefficients (null rejected). $[]^{-}$ ($[]^{+}$) indicate the number of cases that adverse (advantageous) selection will not be rejected based on the joint interpretation of the estimates in the corresponding approach when we rely on statistical significance. $[]^{n-}$ ($[]^{n+}$) indicate the number of cases of adverse (advantageous) selection if we rely solely on the signs of the coefficients.

4.A2 Simulation 2

Scenario	Parameter	Selection	FKS			FMG		RC-Model		
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		$\hat{\alpha}_2$	$\hat{\gamma}_2$	$\mathrm{E}(\hat{\alpha}_2 - \hat{\gamma}_2)$	$\hat{\gamma}_3$	$\hat{\beta}$	$\hat{\delta}$	$\mathrm{E}(\hat{\beta}_i)$	$\mathrm{E}(\hat{\delta}_i)$	$Corr(\hat{\beta}_i, \hat{\delta}_i)$
2a	$\mathrm{E}(\delta_i)$=0, $\mathrm{E}(\beta_i)$=0,	0.080 [100]	0.080 [100]	0.000 [56]	0.000 [53]	0.002 [49]	0.000 [54]	0.002 [49]	0.000 [47]	0.229 [100]
	$\mathrm{Cov}(\beta_\mathrm{i}, \delta_\mathrm{i})$=0.4	[0] [100]	[0] [88]	[0] [0]	[3] [3]	[0] [5]	[3] [3]	[3] [3]	[3] [4]	[0] [100]
	$\sigma_\beta = 0.5, \sigma_\delta = 0.5$			$[56]^{n-}$ $[44]^{n+}$	$[0]^{-}$ $[0]^{+}$	$[0]^{-}$ $[0]^{+}$;	$[57]^{n-}$ $[43]^{n+}$			
2b	$\mathrm{E}(\delta_i)$=0, $\mathrm{E}(\beta_i)$=0,	0.200 [100]	0.199 [100]	0.000 [54]	0.000 [50]	0.003 [48]	0.000 [51]	0.001 [51]	-0.001 [45]	0.324 [100]
	$\mathrm{Cov}(\beta_\mathrm{i}, \delta_\mathrm{i})$=0.4	[0] [100]	[0] [88]	[0] [0]	[3] [3]	[2] [4]	[4] [5]	[3] [1]	[4] [4]	[0] [100]
	$\sigma_\beta = 1, \sigma_\delta = 1$			$[54]^{n-}$ $[46]^{n+}$	$[0]^{-}$ $[0]^{+}$	$[2]^{-}$ $[0]^{+}$;	$[61]^{n-}$ $[39]^{n+}$			
2c	$\mathrm{E}(\delta_i)$=0, $\mathrm{E}(\beta_i)$=0,	0.276 [100]	0.276 [100]	0.000 [56]	0.000 [50]	0.004 [50]	0.000 [48]	0.001 [52]	-0.003 [45]	0.356 [100]
	$\mathrm{Cov}(\beta_\mathrm{i}, \delta_\mathrm{i})$=0.4	[0] [100]	[0] [88]	[0] [0]	[3] [2]	[2] [3]	[3] [6]	[3] [1]	[4] [3]	[0] [100]
	$\sigma_\beta = 1.5, \sigma_\delta = 1.5$			$[56]^{n-}$ $[44]^{n+}$	$[0]^{-}$ $[0]^{+}$	$[2]^{-}$ $[0]^{+}$;	$[62]^{n-}$ $[38]^{n+}$			
Simulations		100	100	100	100	100	100	100	100	100
$N \times T$		1000 × 5	1000 × 5	1000 × 5	1000 × 5	1000 × 5	1000 × 5	1000 × 5	1000 × 5	1000 × 5

Notes: Each scenario reflects the combination of $\mathrm{E}(\delta_i)$ and $\mathrm{E}(\beta_i)$ with a correlation structure $\mathrm{Cov}(\beta_\mathrm{i}, \delta_\mathrm{i})$, σ_β and σ_δ. Column (1) indicates whether there is a significant correlation between insurance (I) and risk (R), indicating IA. Columns (2) to (4) show the estimates of the structural equations, (4.5) and (4.6), of the FKS approach. Column (3) can be interpreted as the degree of IA associated with z (the 'omitted variable bias') suggested by the FKS model: An F-test is used to test whether the coefficient of R (based on the specification in column (2)) equals the coefficient of R in column (1). Columns (5) and (6) are based on estimates of the equations (4.3) and (4.4) from the FMG approach. Column (5) provides the estimate of the effect of z on R in the FMG approach, and column (6) shows its effect on I. These estimates are based on least squares estimates. Columns (7) and (8) show coefficients from a multilevel model where parameters δ and β are estimated for each individual. Standard errors are clustered at the individual level. Column (9) shows the correlation between $\hat{\beta}_i$ and $\hat{\delta}_i$. [] indicate counts from 1 to 100. [] on the right side of coefficients counts the number of positive values. [] on the left hand side under a coefficient counts the number of cases a coefficient is both negative and significantly different from zero (5 % level); [] on the right side under a coefficient suggests the number of positive coefficients (null rejected). $[]^{-}$ ($[]^{+}$) indicate the number of cases that adverse (advantageous) selection will not be rejected based on the joint interpretation of the estimates in the corresponding approach when we rely on statistical significance. $[]^{n-}$ ($[]^{n+}$) indicate the number of cases of adverse (advantageous) selection if we rely solely on the signs of the coefficients.

4.A3 Simulation 3

Scenario	Parameter	Selection	FKS	FMG				RC-Model		
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		$\hat{\alpha}_2$	$\hat{\gamma}_2$	$E(\hat{\alpha}_2 - \hat{\gamma}_2)$	$\hat{\gamma}_3$	$\hat{\beta}$	$\hat{\delta}$	$E(\hat{\beta}_i)$	$E(\hat{\delta}_i)$	$Corr(\hat{\beta}_i, \hat{\delta}_i)$
3a	$E(\delta_i)$=-0.5, $E(\beta_i)$=0.3, $Cov(\beta_i, \delta_i)$=0	-0.073 [0] [84] [0]	0.000 [47] [0] [0]	-0.072 [0] [98] [0] $[0]^{n-}$ $[100]^{n+}$	-0.500 [0] [100] [0] $[0]^{-}$ $[98]^{+}$	0.303 [100] [0] [100] $[0]^{-}$ $[100]^{+}$;	-0.500 [0] [100] [0] $[0]^{n-}$ $[100]^{n+}$	0.301 [100] [0] [100]	-0.501 [0] [100] [0]	-0.001 [47] [2] [2]
3b	$E(\delta_i)$=-0.5, $E(\beta_i)$=0.3, $Cov(\beta_i, \delta_i)$=0.3	0.071 [100] [0] [84]	0.149 [100] [0] [88]	-0.079 [0] [93] [0] $[0]^{n-}$ $[100]^{n+}$	-0.545 [0] [100] [0] $[0]^{-}$ $[93]^{+}$	0.303 [100] [0] [100] $[0]^{-}$ $[100]^{+}$;	-0.500 [0] [100] [0] $[0]^{n-}$ $[100]^{n+}$	0.301 [100] [0] [100]	-0.501 [0] [100] [0]	0.242 [100] [0] [0] [100]
3c	$E(\delta_i)$=-0.5, $E(\beta_i)$=0.3, $Cov(\beta_i, \delta_i)$=0.6	0.214 [100] [0] [100]	0.299 [100] [0] [88]	-0.085 [0] [68] [0] $[0]^{n-}$ $[100]^{n+}$	-0.590 [0] [100] [0] $[0]^{-}$ $[68]^{+}$	0.303 [100] [0] [100] $[0]^{-}$ $[100]^{+}$;	-0.499 [0] [100] [0] $[0]^{n-}$ $[100]^{n+}$	0.301 [100] [0] [100]	-0.501 [0] [100] [0]	0.487 [100] [0] [0] [100]
Simulations		100	100	100	100	100	100	100	100	100
$N \times T$		1000 × 5	1000 × 5	1000 × 5	1000 × 5	1000 × 5	1000 × 5	1000 × 5	1000 × 5	1000 × 5

Notes: Each scenario reflects the combination of $E(\delta_i)$ and $E(\beta_i)$ with the correlation structure $Cov(\beta_i, \delta_i)$, σ_β and σ_δ set to 1. Column (1) indicates whether there is a significant correlation between insurance (I) and risk (R), indicating IA. Columns (2) to (4) show estimates of the structural equations, (4.5) and (4.6), of the FKS approach. Column (3) can be interpreted as the degree of IA associated with z (the 'omitted variable bias') suggested by the FKS model: An F-test is used to test whether the coefficient of R (based on the specification in column (2)) equals the coefficient of R in column (1). Columns (5) and (6) are based on estimates of the equations (4.3) and (4.4) from the FMG approach. Column (5) provides the estimate of the effect of z on R in the FMG approach, and column (6) shows its effect on I. These estimates are based on least squares estimates. Columns (7) and (8) show coefficients from a multilevel model where parameters δ and β are estimated for each individual. Standard errors are clustered at the individual level. Column (9) shows the correlation between $\hat{\beta}_i$ and $\hat{\delta}_i$. [] indicate counts from 1 to 100. [] on the right side of coefficients counts the number of positive values. [] on the left hand side under a coefficient counts the number of cases a coefficient is both negative and significantly different from zero (5 % level); [] on the right side under a coefficient suggests the number of positive coefficients (null rejected). $[]^{-}$ ($[]^{+}$) indicate the number of cases that adverse (advantageous) selection will not be rejected based on the joint interpretation of the estimates in the corresponding approach when we rely on statistical significance. $[]^{n-}$ ($[]^{n+}$) indicate the number of cases of adverse (advantageous) selection if we rely solely on the signs of the coefficients.

4.A4 Simulation 4

Scenario	Parameter	Selection (1) $\hat{\alpha}_2$	FKS (2) $\hat{\gamma}_2$	(3) $\mathrm{E}(\hat{\alpha}_2 - \hat{\gamma}_2)$	(4) $\hat{\gamma}_3$	FMG (5) $\hat{\beta}$	(6) $\hat{\delta}$	RC-Model (7) $\mathrm{E}(\hat{\beta}_i)$	(8) $\mathrm{E}(\hat{\delta}_i)$	(9) $Corr(\hat{\beta}_i, \hat{\delta}_i)$
4a	$\mathrm{E}(\delta_i)$=0.3, $\mathrm{E}(\beta_i)$=0.4, $\mathrm{Cov}(\beta_\mathrm{i}, \delta_\mathrm{i})$=0	0.056 [100] [0] [70]	0.000 [47] [0] [0]	0.056 [100] [0] [91] [100]$^{n-}$ [0]$^{n+}$	0.300 [100] [0] [100] [91]$^{-}$ [0]$^{+}$	0.403 [100] [0] [100] [100]$^{-}$ [0]$^{+}$;	0.300 [100] [0] [100] [100]$^{n-}$ [0]$^{n+}$	0.401 [100] [0] [100]	0.299 [100] [0] [100]	-0.001 [47] [2] [2]
4b	$\mathrm{E}(\delta_i)$=0.3, $\mathrm{E}(\beta_i)$=0.4, $\mathrm{Cov}(\beta_\mathrm{i}, \delta_\mathrm{i})$=-0.3	-0.083 [0] [96] [0]	-0.150 [0] [88] [0]	0.067 [100] [0] [100] [100]$^{n-}$ [0]$^{n+}$	0.360 [100] [0] [100] [100]$^{-}$ [0]$^{+}$	0.403 [100] [0] [100] [100]$^{-}$ [0]$^{+}$;	0.299 [100] [0] [100] [100]$^{n-}$ [0]$^{n+}$	0.401 [100] [0] [100]	0.299 [100] [0] [100]	-0.245 [0] [100] [0]
4c	$\mathrm{E}(\delta_i)$=0.3, $\mathrm{E}(\beta_i)$=0.4, $\mathrm{Cov}(\beta_\mathrm{i}, \delta_\mathrm{i})$=-0.6	-0.222 [0] [100] [0]	-0.300 [0] [88] [0]	0.078 [100] [0] [92] [100]$^{n-}$ [0]$^{n+}$	0.420 [100] [0] [100] [92]$^{-}$ [0]$^{+}$	0.403 [100] [0] [100] [100]$^{-}$ [0]$^{+}$;	0.299 [100] [0] [100] [100]$^{n-}$ [0]$^{n+}$	0.401 [100] [0] [100]	0.299 [100] [0] [100]	-0.487 [0] [100] [0]
Simulations		100	100	100	100	100	100	100	100	100
$N \times T$		1000 × 5	1000 × 5	1000 × 5	1000 × 5	1000 × 5	1000 × 5	1000 × 5	1000 × 5	1000 × 5

Notes: Each scenario reflects the combination of $\mathrm{E}(\delta_i)$ and $\mathrm{E}(\beta_i)$ with the correlation structure $\mathrm{Cov}(\beta_\mathrm{i}, \delta_\mathrm{i})$, σ_β and σ_δ set to 1. Column (1) indicates whether there is a significant correlation between insurance (I) and risk (R), indicating IA. Columns (2) to (4) show estimates of the structural equations, (4.5) and (4.6), of the FKS approach. Column (3) can be interpreted as the degree of IA associated with z (the 'omitted variable bias') suggested by the FKS model: An F-test is used to test whether the coefficient of R (based on the specification in column (2)) equals the coefficient of R in column (1). Columns (5) and (6) are based on estimates of the equations (4.3) and (4.4) from the FMG approach. Column (5) provides the estimate of the effect of z on R in the FMG approach, and column (6) shows its effect on I. These estimates are based on least squares estimates. Columns (7) and (8) show coefficients from a multilevel model where parameters δ and β are estimated for each individual. Standard errors are clustered at the individual level. Column (9) shows the correlation between $\hat{\beta}_i$ and $\hat{\delta}_i$. [] indicate counts from 1 to 100. [] on the right side of coefficients counts the number of positive values. [] on the left hand side under a coefficient counts the number of cases a coefficient is both negative and significantly different from zero (5 % level); [] on the right side under a coefficient suggests the number of positive coefficients (null rejected). []$^{-}$ ([]$^{+}$) indicate the number of cases that adverse (advantageous) selection will not be rejected based on the joint interpretation of the estimates in the corresponding approach when we rely on statistical significance. []$^{n-}$ ([]$^{n+}$) indicate the number of cases of adverse (advantageous) selection if we rely solely on the signs of the coefficients.

4.A5 Data description

Variable	Description
PHI	owner of private health insurance
HR	2 lowest health categories based on self assessed health
smoke_now	current smoker
smoke_past	past smoker
female	women
age	actual age of respondent
working	respondent is employed or self employed
not_work	is not at the job market
retired	respondent is retired
educ	more than no qualification
couple	married or cohabit
nwhite	ethnicity recoded to non white
children	respondent has children
ill1	diagnosed cardiovascular diseases
ill2	pneumological disease, asthma
ill3	arthritis, osteoporosis
ill4	cancer
ill5	psychiatric disorder, Alzheimer's
ill6	specific eye problems

4.A6 Estimates standard approaches

	(1)	(2)	(3)	(4)
Approach	FMG		FKS	
Dep. var.	privins	HR	privins	privins
HR			-0.034***	-0.035***
			(0.000)	(0.000)
time_avail	-0.014**	-0.076***		-0.017***
	(0.026)	(0.000)		(0.008)
female	0.020***	-0.043***	0.018**	0.018**
	(0.006)	(0.000)	(0.011)	(0.011)
age	-0.001**	0.002***	-0.001**	-0.001**
	(0.015)	(0.000)	(0.017)	(0.022)
black	-0.031	0.187***	-0.024	-0.024
	(0.266)	(0.000)	(0.392)	(0.381)
couple	0.037***	-0.064***	0.035***	0.035***
	(0.000)	(0.000)	(0.000)	(0.000)
children	-0.033***	0.030**	-0.033***	-0.032***
	(0.003)	(0.018)	(0.004)	(0.005)
educ	0.072***	-0.107***	0.069***	0.068***
	(0.000)	(0.000)	(0.000)	(0.000)
working	0.011	-0.132***	0.009	0.007
	(0.185)	(0.000)	(0.271)	(0.441)
not_work	0.006	0.152***	0.012	0.011
	(0.446)	(0.000)	(0.123)	(0.148)
smoke_now	-0.053***	0.162***	-0.048***	-0.048***
	(0.000)	(0.000)	(0.000)	(0.000)
smoke_past	-0.012	0.053***	-0.010	-0.010
	(0.120)	(0.000)	(0.192)	(0.188)
_cons	0.087**	0.362***	0.087**	0.100***
	(0.014)	(0.000)	(0.013)	(0.005)
N	31431	31431	31431	31431

Notes: All columns show coefficients from a linear probability model. Columns 1 and 2 are estimates from the FMG approach, whereas columns 3 and 4 reflect the coefficients based on FKS. Regional dummies are included and standard errors are clustered at individual level; p-values in parentheses. * p <0.10, ** p <0.05, *** p <0.01

Chapter 5

The Relationship Between Public Health Insurance and Informal Transfer Networks in Ghana[1]

5.1 Introduction

In the developing world, individual access to health services is largely determined by income. The 'cash and carry' system that is prevalent in most developing countries restricts medical access to the amount of money directly paid to healthcare providers. In order to be able to afford treatment costs, many poor households rely on informal transfers within networks of relatives or neighbours. These support schemes are important and beneficial, since the risk of becoming sick can be shared with other members of the network (Fafchamps, 2008). An individual's engagement in an informal transfer network is usually governed by two main motives: altruism and reciprocity (Leider et al., 2009). Altruism is a preference for contributing without expectations of being rewarded, while reciprocity is based on an exchange motive with the prospect of future benefits (e.g., Cox, 1987; Ligon and Schechter, 2012). Support schemes can be a crucial insurance mechanism in times of severe hardship in Ghana (Tsai and Dzorgbo, 2012) and frequently support individuals during the more vulnerable stages of their lives, e.g., when they are young or very old (Kabki, 2007).

However, these networks can provide inadequate protection if many members are suffering from economic hardship or refuse to contribute due to personal conflicts

[1]This study is joint work with Christoph Strupat. See Klohn and Strupat (2013) for an older working paper.

(Townsend, 1994; Morduch, 1999). Kinship networks in particular are often characterized by strong sharing obligations, meaning productive network members face demands for transfers from less productive relatives (Platteau, 2000; Hoff and Sen, 2005; Di Falco and Bulte, 2011). This redistributive pressure can adversely affect the incentives for enterprise-owning network members to invest in their own businesses (Grimm et al., 2013) or save above a certain amount (Duflo et al., 2009; Wahhaj, 2010; Brune et al., 2015). Thus, adverse incentives prevent members from improving their economic situations and may be an important barrier to economic transition.

To overcome the imperfections of informal transfer networks and help relatively productive individuals such as enterprise owners develop their full economic potential, formal health insurance schemes or micro-insurance are seen as an important remedy (Landmann et al., 2012). In recent years, some developing countries (India, Ghana and Nigeria) have introduced country-wide health insurance schemes, while in other countries many micro-insurance initiatives have been launched to complement informal insurance mechanisms. While there is already some empirical evidence for the crowding out of informal mechanisms after receiving public transfers (Dercon and Krishnan, 2003; Pavan and Colussi, 2008; Oruč et al., 2011), there have been few studies into the relationship between formal insurance and informal transfer networks. None of these studies have investigated the effect of a formal, nationwide health insurance scheme. Attanasio and Rıos-Rull (2000) provide theoretical and empirical evidence that formal insurance crowds out informal insurance and potentially increases welfare in Mexico. Dubois et al. (2008) analyse the interaction between formal and informal agreements using data from Pakistan. They assert that policy interventions can complement or weaken informal cooperation. Landmann et al. (2012) ran an experiment in the rural Philippines and showed that formal insurance can lead to lower voluntary transfers among network members. In a theoretical and experimental analysis, Lin et al. (2014) found that introducing formal insurance significantly crowds out private transfers and reduces income inequality.

To the best of our knowledge, this chapter provides the first empirical evidence on whether informal transfers are affected by a formal, nationwide health insurance scheme. The launch of the Ghanaian National Health Insurance Scheme (NHIS) in 2003, coupled with differences in the date of implementation between local districts, makes Ghana an ideal setting for examining the relationship between formal health insurance and informal transfer networks. We also examine the impact of the NHIS on health-related outcomes such as health status and out-of-pocket (OOP) payments, which may help understand how the implementation of the NHIS has affected infor-

mal transfers. Our study contributes to the literature in three ways. First, while many studies on this topic use experimental methods that may have limited external validity, we use survey data which is representative for the entire population. Second, we look at the exogenous introduction of a public health insurance scheme using a quasi-experimental setup that allows us to evaluate the causal impact of a formal health insurance scheme on informal transfer behaviour. Third, as sharing obligations are generally strong within kinship networks in Ghana (Udry and Conley, 2004), we use detailed information on the relationship status between donors and recipients of transfers (e.g., parents, siblings, non-relatives) to explore whether the impact of the NHIS varies with relationship status.

As the health insurance scheme was implemented on different dates by most district authorities between 2005 and 2006, we use the fifth wave of the Ghanaian Living Standard Household Survey (GLSS), which was conducted over a 12-month survey period (October 2005 to September 2006). The districts in this cross-sectional household survey contain enumeration areas (which we call sub-districts in the following) that were interviewed in different months during the survey period. We use this variation in interview dates for our identification strategy. In particular, we are able to identify the sub-districts that were interviewed before and after the implementation of the NHIS, as we use the exact implementation dates of the NHIS, which vary at the district level. In addition, we also identify those districts that implemented the NHIS after the survey period. Thus, our identification strategy uses a difference-in-difference framework comparing individuals at different points in time (interview months) who live in districts where the NHIS has been implemented with individuals where it has not.

In our empirical analysis, we first estimate a linear probability model (LPM) to evaluate the extent to which the implementation of the NHIS influenced the probability of sending or receiving regular transfers[2]. In a second step, we investigate the extent to which the NHIS affected the number of transfers sent and received. Our econometric specifications control for district-specific unobserved characteristics (such as supply-side healthcare provision factors) and seasonality during the course of the year. We also examine the impact of the NHIS on health-related outcomes such as the probability of low health status (reflecting whether a respondent had to stop usual activities for two weeks), the number of sick days during the previous two weeks and OOP expenditures.

[2]We use variables that show whether household members send and receive transfers in the form of money or goods on a weekly, monthly or quarterly basis within Ghana.

We find that introducing a formal health insurance scheme has no effect on health, but substantially reduces OOP expenditures, which in turn lowers the need for regular informal transfers for health purposes. As a consequence, we find a reduction in remittances to other households for health purposes. In particular, the amount of remittances to non-relatives is reduced, which might be due to relatively low sharing obligations between unrelated network members. As the NHIS covers all outpatient/inpatient services and also treatment for chronic diseases, our results show that it is not only ill individuals who benefit financially from the NHIS, but also donors who are relieved of their financial burden.

The remainder of this chapter is organized as follows. Section 2 introduces the theoretical framework of our study and provides information about the national health insurance scheme in Ghana. In section 3 we describe the data and provide details about our identification strategy. Section 4 presents the results and further robustness checks. Finally, section 5 concludes with a summary of the main findings and a research outlook.

5.2 Theoretical Framework and the National Health Insurance Scheme in Ghana

5.2.1 Theoretical Framework

In Ghana, reciprocity is widespread and often necessary to reduce economic insecurity, building trust and solidarity within transfer networks (Udry and Conley, 2004). From an economic perspective, reciprocity is an exchange motive with respect to future benefits (e.g., Cox, 1987; Ligon and Schechter, 2012) which drives the formation of transfer networks as an informal institution and provides signals for being trustworthy which can foster an individual's social status. In Ghana, transfer networks are largely made up of relatives, forming kinship networks in which reciprocal transfers are used to generate responsibility and obligations between network members. These networks can provide a crucial insurance mechanism in times of severe hardship (Tsai and Dzorgbo, 2012), but also regularly support individuals during the more vulnerable stages of their lives, e.g., when they are young or very old (Kabki, 2007). Thus, reciprocity is an important driver for participating in informal transfer networks in Ghana, either through direct financial benefits due to risk sharing, or through indirect benefits such as increased social status within the community.

Against this background, our theoretical framework assumes that an individual's

engagement in an informal transfer network is determined by the exchange motive (reciprocity). In line with Morduch (1999), the framework is based on two individuals who form a transfer network with repeated interactions over time. Both individuals contribute to the network until one individual reneges on the arrangement. Hence, there is a trade-off between leaving the network today and future benefits from further participation. A rational individual will make a cost-benefit analysis which consists of three components: (future) benefits of the reciprocal arrangement in terms of received transfers, costs as determined by (current) transfer payments to the network partner and finally exit costs for leaving the network, as determined by the relationship between network partners and their respective sharing obligations. This analysis also contains opportunity costs due to a possible (partial) substitute (e.g., a formal insurance scheme). Including exit costs is in line with the social network analysis literature, which characterises networks between immediate family members as generally closed and associated with strong sharing obligations and high (psychological) exit costs (Granovetter, 1973; Grimm et al., 2013). In contrast, networks between non-relatives are often characterised by lower sharing obligations and exit costs.

The decision to stay in the network is negatively correlated with opportunity costs and (current) transfer payments. In contrast, it is positively correlated with expected benefits and exit costs. If, for example, transfer payments to the network partner are low and the expected benefits of this reciprocal arrangement are high, the individual will stay within the network. After the introduction of a formal insurance scheme, new opportunity costs can change this decision if overall costs (including opportunity costs) exceed expected benefits. Thus, an individual with higher opportunity costs may decide to leave the network or reduce transfer payments. Importantly, high (low) exit costs will decrease (increase) the probability of reducing the contribution to the network.

Two predictions arise from our simple theoretical framework. First, we expect that the availability of the NHIS will decrease informal network participation at the extensive and intensive margin, for both sending and receiving transfers. Second, we expect that the degree of crowding out will depend strongly on the relationship status and sharing obligations between the network partners, i.e., the closer the relationship (strong kinship ties and high exit costs), the lower the crowding out.

5.2.2 The National Health Insurance Scheme in Ghana

The Ghanaian Parliament passed the National Health Insurance Scheme into law in 2003. The scheme was incrementally implemented at the district level by the end of 2006. The aim of the scheme is to provide healthcare services to a broad swathe of the population and to establish an alternative to the existing 'cash and carry' system. The insurance covers all basic outpatient, inpatient and dental health services such as X-rays, blood tests, malaria treatments, surgical operations and maternal care services (e.g., antenatal care, deliveries and postnatal care). In this latter category, the NHIS has been effective in improving health outcomes for recent mothers, who are now more likely to receive prenatal care, deliver at a hospital, have their deliveries attended by trained health professionals and experience fewer birth complications (Mensah et al., 2010).

Membership in the health insurance scheme is voluntary for all adults (age 14-69) who work in the informal sector, such as self-employed individuals, while membership is mandatory for formal sector employees, with insurance premiums deducted from their monthly payrolls. The income-related insurance premium varies between a minimum of 7.2 Ghanaian cedis (GHC) (US$3) and a maximum of 48.0 GHC (US$19), paid on an annual basis.[3] All children under 14 whose parents have enrolled in the scheme and all people over 69 are covered by the insurance but exempted from paying premiums.

The NHIS is monitored and regulated by the National Health Insurance Authority (NHIA). Covered health services are mainly financed by a health insurance levy (a 2.5% addition to the value added tax), payment of insurance premiums and money allocated by the government. The NHIA licenses district mutual health insurance schemes (DMHISs) that are established by the district authorities to collect sufficient insurance premiums to meet the expected healthcare claims within each district. After a DMHIS has paid two million GHC into the NHIA, and health insurance cards have been distributed to inhabitants who paid the insurance premium, the DMHIS is officially launched and all basic healthcare services are covered by the insurance (Seddoh et al., 2011). As acceptance of health insurance and the financial ability of each district varies, the DMHISs were implemented at different dates, with most district authorities launching the scheme in 2005 and 2006.

[3] 1GHC=0.4US$

5.3 Data and Identification Strategy

5.3.1 Data Description

Our analysis uses the fifth wave of the Ghana Living Standards Survey (GLSS5), which is based on interviews conducted by the Ghana Statistical Office and the World Bank during the period between October 2005 and September 2006. This nationwide survey contains socio-economic variables measured at the individual and household levels, including information on informal transfer networks. It is the source for a nationally representative sample of 8,687 households living in 110 districts and 580 sub-districts, with 37,128 household members.

Our treatment variable is a binary indicator representing the availability of the NHIS in an individual's district. In order to construct this variable, we collected the exact implementation dates of the NHIS at the district level by contacting district officials and using district-specific media reports on health insurance. Figure 5.1 shows how the NHIS implementation evolved over time and districts.[4] The two dashed lines indicate the start and end of the survey period.

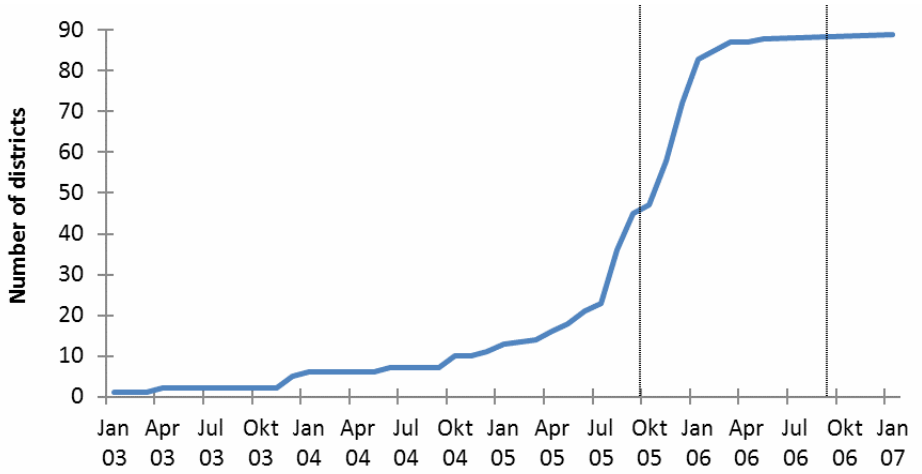


Figure 5.1 NHIS availability at the district level

As measures for participating in reciprocal transfer networks, we define variables that show whether household members sent or received transfers in the form of money or goods on a weekly, monthly or quarterly basis within Ghana. In addition, we also use information on the number of transfers made and received. These regular transfers include no labour compensation for extended family members or

[4]We use 90 out of 110 districts for our analysis, as the district authorities provided the exact date of the NHIS implementation.

neighbours who work in a business linked to the household. Most of these transfers occurred within inner family networks, especially to children/parents (50 percent) and extended family members such as grandparents, cousins and aunts and uncles (15 percent), while transfers to non-relatives were less common (35 percent). Although we analyse the short-term effects of the policy, we combine all monetary values into an annual amount to simplify comparisons with other financial information that is provided on an annual basis. As 42 percent of all household members did not provide information for both transfer variables, we investigate the impact of the NHIS implementation on sent and received transfers separately.[5] Thus, our analysis focuses on all individuals who are not exempted from premium payments. 5,956 individuals living in 2,710 households gave information on regularly sent transfers, while information on regularly received transfers was available for 4,985 individuals in 2,611 households.

5.3.2 Identification Strategy

To investigate the relationship between informal transfer networks and formal health insurance, our identification strategy is based on a quasi-experimental setup. We collected data on the precise implementation dates for the NHIS at the district level, i.e., when health insurance coverage became available, and benefit from the fact that the district's sub-districts were surveyed at different points during the survey period between October 2005 and September 2006. As most districts introduced the NHIS during this survey period, we were able to use the variation in interview dates to compare individuals who were interviewed before and after the introduction of the insurance scheme. However, some districts did not implement the NHIS during the survey period or were entirely surveyed before it was implemented. For instance, the Nkwanta district in the Volta region is divided into eight sub-districts, four of which were surveyed in November 2005, with the other four interviewed in March 2006. The Nkwanta district introduced the NHIS in January 2006. In comparison, the Nanumba district also consists of eight sub-districts and was surveyed during the same months, but the NHIS was not implemented until July 2006.

Thus, in our identification strategy we compare individuals at different points in time (interview months) that are living in districts where the NHIS is implemented

[5]We also have information on the purpose of the transfers, obtained via the question: 'Please rank the three main uses of sent/received transfers?'. 58 percent indicated 'health' as one of the two main reasons for making transfers. However, as this question is entirely self-assessed and does not allow us to make quantitative statements, we stick with the general indicator.

and where it is not, once time- and district-fixed effects are partialled out. In order to control for time-invariant district characteristics such as financial ability or health infrastructure that are likely to be correlated with both the timing of NHIS implementation and our dependent variables, we include district dummies into all our specifications. We also include interview month dummies in order to allow for changes in the macroeconomic situation during the course of the year, which likely affect individuals in the treatment and control group similarly. More formally, our estimates are based on the following equation:

$$y_{idt} = \beta_0 + \beta_1 NHIS_{idt} + \mu_d + \delta_t + \epsilon_{idt} \tag{5.1}$$

The dependent variable y_{idt} indicates if respondent i that lives in district d and was surveyed in month t, makes (receives) transfers. This variable is regressed on the binary treatment variable $NHIS_{idt}$, which takes the value 1 if the respondent was surveyed after the district implemented the NHIS and 0 otherwise. β_0 is a constant, while μ_d represents district fixed effects and δ_t interview month fixed effects. The results we provide are based on variation which is orthogonal to the district- and time-specific part of our specification's error term. In accordance with the concept of potential outcomes (Rubin, 1974, 1977), our identification strategy is based on the assumption:

$$\mathbb{E}(y_{0idt}|d,t) = \mu_d + \delta_t \tag{5.2}$$

i.e., if the NHIS had not been implemented, the potential outcomes, y_{0idt}, would solely depend on district-specific levels, μ_d, and common time effects, δ_t (cf. Angrist and Pischke, 2009). As discussed above, in our application the length of time periods before and after the introduction depends on the time of the interview and the date when the insurance scheme was introduced. In order to increase the precision of our estimates and to control for confounding factors that might be correlated with the introduction of the NHIS and the dependent variable, we furthermore include individual and household specific variables X_{idt} in our specifications. Such variables reflect important socio-demographic differences but also indicate if the respondent is living in an urban or rural sub-district.[6] Thus, we extend equation (5.1):

$$y_{idt} = \beta_0 + \beta_1 NHIS_{idt} + \beta_2^{'} X_{idt} + \mu_d + \delta_t + \epsilon_{idt} \tag{5.3}$$

[6]The description of the variables is presented in the next section.

Our coefficient of interest is β_1, which represents an intention-to-treat effect (ITT) i.e. the effect of an offer to participate in the NHIS on the individual's transfer behavior. This parameter has a causal interpretation, if no additional differences between the treatment and control group exist that can be traced back to the introduction of the NHIS. This assumption is not directly testable. As urban sub districts are probably differently affected by changes in the economic situation, we supplement our identification strategy by interacting the interview month with the urban ecological area dummy allowing for an additional time trend and to evaluate whether our findings are triggered by urban specific changes over time.

Our use of the variation in NHIS implementation dates across districts and the variation in interview dates within districts has two potential sources of bias which may confound our estimate of β_1. First, the order of the NHIS rollout might have been driven by time-invariant district characteristics that are also correlated with our dependent variables. For example, wealthy districts that spend more than average on transfers might have been able to implement the NHIS earlier than less affluent districts. In order to check whether our outcome variables are fundamentally different for districts with and without the NHIS, we provide a balance table of our dependent variables using the fourth wave of the GLSS (1998/1999)[7].

Second, interview dates may have been driven by heterogeneity between sub-districts that also influence the potential outcomes of our analysis. If, for example, the sub-districts were not randomly surveyed over time and the survey team interviewed urban sub-districts first, this would bias our estimates for NHIS implementation. To investigate the extent to which observed changes in NHIS implementation were triggered by structural heterogeneity among sub-districts, we conduct several estimates as robustness checks in section 5.4. We conduct regressions using predetermined, time-invariant characteristics of the sub-districts and their inhabitants such as education, gender and an indicator of whether the respondent was living in an urban or rural sub-district as dependent variables. We also conduct placebo regressions by using the fourth wave of the GLSS (1998/1999). Thus, we can gauge whether a systematic relationship between the sub-districts interviewed at different times and the dependent variable would bias our estimate of β_1.

[7]This survey was conducted in the same manner over a 12-month period between 1998 and 1999 and contains the same number of districts as the GLSS from 2005/2006.

5.3.3 Estimation Models

We begin by generating a linear probability model (LPM) to evaluate whether the introduction of the NHIS influenced the probability that a household engaged in transfers. Our dependent variable is a dummy variable which takes the value 1 if the respondent transferred money or goods to non-household members and 0 if no transfers took place. We employ the same model to investigate whether an individual received transfers. The binary nature of the dependent variable would conventionally suggest the estimation of a probit or logit model. Binary choice models, however, can be problematic when applied using the least-squares dummy variable approach, because they suffer from the incidental parameters problem and a substantial loss of observations. In a second step, we examine the extent to which the number of sent or received transfers was affected by the implementation of the NHIS. Therefore, we estimate a regression model with the actual value of either sent or received transfers as the dependent variable.

5.4 Results and Robustness Checks

Before we turn to the empirical results, Table 5.1 displays the means of our dependent variables, distinguished by the availability of the NHIS. Less money was sent to other households by respondents living in areas with the NHIS available. Among the respondents able to use the NHIS, 41% sent money regularly, compared to 70% of individuals living in areas without the NHIS. In addition, the average transfer amount was 44 GHC less among respondents from areas without the NHIS, 38 % less than in areas with the NHIS. To understand how the implementation of the NHIS affected informal transfers, we also examine the impact of the NHIS on health-related outcomes such as the probability of low health status (in which a respondent had to stop their usual daily activities for two weeks or more), the number of sick days taken during the previous two weeks and OOP expenditures. All indicators are lower for sub-districts with the NHIS available.

To check whether districts with and without the NHIS were fundamentally different before the implementation of the NHIS, we provide a balance table that shows the means of our dependent variables from the fourth wave of the GLSS, conducted in 1998/1999 (see Table 5.2). As can be seen from the p-values, two-sided tests comparing the values for the two groups do not show statistically significant differences. This indicates that both groups were balanced across all outcome variables before the NHIS implementation.

Table 5.1 Descriptive statistics

	(1) Total mean	(2) NHIS mean	(3) No NHIS mean	(4) Difference in means	(5) p-value
Sent transfers (0/1)	0.46	0.41	0.69	-0.28	0.00
Value of sent transfers	75.77	68.49	111.95	-43.46	0.00
N	5956	4955	1001		
Received transfers (0/1)	0.35	0.32	0.47	-0.15	0 .00
Value of received transfers	54.52	53.86	59.39	-5.53	0.38
N	4985	4054	931		
Low health status	0.13	0.12	0.19	-0.07	0.00
Number of sick days	0.47	0.77	1.25	-0.48	0.00
Medical expenditures	33.67	31.5	45.12	-13.62	0.00
N	5009	4025	984		

Notes: For our calculation we use the GLSS (2005/2006).

Table 5.2 Difference in means before the implementation of the NHIS (GLSS 1998/1999)

	(1) NHIS mean	(2) No NHIS mean	(3) Difference in means	(4) p-value
Sent transfers (1/0)	0.26	0.24	0.02	0.16
Value of sent transfers	62.12	55.02	7.1	0.28
N	4807	934		
Received transfers (1/0)	0.16	0.12	0.03	0.35
Value of received transfers	49.12	45.83	3.29	0.45
N	5003	926		
Low health status (1/0)	0.18	0.18	0.00	0.86
Number of sick days	1.07	1.21	-0.14	0.27
Out-of-pocket expenditures	85.81	74.94	10.87	0.46
N	4807	934		

Notes: Our calculations use the 1998/1999 GLSS, which was conducted over a 12-month period between 1998 and 1999 in the same manner and for the same number of districts as the 2005/2006 GLSS.

Table 5.3 provides estimates of the NHIS implementation on health-related outcomes. We find negative effects of the NHIS implementation for all outcomes. Having the NHIS available improved health status and reduced the average number of sick days by 0.38, but the coefficients are not statistically significant. OOP expenditures decreased by an average of 26 GHC, a relative reduction of 58 %. Our findings are in line with Powell-Jackson et al. (2014), who examine the impact of removing user fees for healthcare using data from a randomized control trial in Ghana. The authors also find a reduction in OOP expenditures and no statistically significant effect on health. These results suggest that sick individuals were financially relieved by the implementation of the NHIS. To examine whether transfer behaviour was also affected, the first three columns in Table 5.4 present estimates from the LPM using sent transfers as the dependent variable.

Table 5.3 Effect of NHIS implementation on health status, number of sick days and medical expenditures

Dependent Variables:	LPM Low health	OLS 1 # of sick days	OLS 2 OOP expend.
NHIS	-0.047 (0.034)	-0.382 (0.261)	-26.320*** (8.578)
N	5009	5009	5009
adj. R-sq	0.030	0.041	0.027
District and interview month dummies	Yes	Yes	Yes
Individual and HH control variables	Yes	Yes	Yes
Urban time trend	Yes	Yes	Yes

Notes: Standard errors (in parentheses) are clustered at the district level. * p <0.10, ** p <0.05, *** p <0.01

The first column shows the NHIS coefficient without including individual and household variables into the estimation model. We find a negative and statistically significant effect for the NHIS dummy. The implementation of the NHIS decreased the probability of transferring money to other households by 15 %. The size of the coefficient remains similar if we include our control variables, which suggests that the implementation of the NHIS is randomly assigned in terms of individual and household specific variables. In addition, including control variables increases the precision of our estimates, as the NHIS coefficient then becomes significant at the 5 % level.

We consider a range of variables that are typically used to control for socioeconomic characteristics such as education level, employment status, age and sex (see Table in Appendix 5.A2). In addition, we include household expenditures as

an important control variable for a household's financial potential (Deaton, 1997). Expenditures are corrected via a region-specific consumer price index and an equivalence scale to reflect age- or sex-specific relative consumption needs (Service, 2008). We also include a dummy that indicates whether a respondent lives in an urban subdistrict. We condition on variables that possibly determine the degree of informal risk sharing. These are household size, marital status, owning a savings account and migration status. We also include an urban-specific time trend (column 3) which does not affect the size of the NHIS coefficient, indicating that our findings are not confounded by regional changes during the survey period.

The last three columns of Table 5.4 contain the estimates for the value of sent transfers as the dependent variable. The implementation of the NHIS led to a crowding out of 24 GHC in the specification with control variables. Interestingly, the reduction in OOP payments (26 GHC, see Table 5.3) is very similar to the reduction in sent transfers (24 GHC), which suggests that sick individuals and donors were financially relieved by the implementation of the NHIS.

Table 5.4 Effect of the NHIS implementation on sending transfers

	LPM 1	LPM 2	LPM 3	OLS 1	OLS 2	OLS 3
NHIS	-0.148*	-0.146**	-0.148**	-23.803	-24.172*	-23.723*
	(0.079)	(0.061)	(0.060)	(14.866)	(14.130)	(14.005)
N	5956	5956	5956	5956	5956	5956
adj. R-sq	0.061	0.163	0.164	0.023	0.150	0.150
District and interview month dummies	Yes	Yes	Yes	Yes	Yes	Yes
Individual and HH control variables	No	Yes	Yes	No	Yes	Yes
Urban time trend	No	No	Yes	No	No	Yes

Notes: Standard errors (in parentheses) are clustered at the district level. * p <0.10, ** p <0.05, *** p <0.01; See Appendix 5.A3 for detailed regression outputs.

To determine whether there is also a reduction in received transfers, we use our estimates with received transfers as the dependent variable. The coefficients indicate a negative but statistically insignificant relationship between the implementation of the NHIS and receiving transfers for all models (see Table 5.5). Looking at the estimates for the value of received transfers as our dependent variable shows the size of the coefficients is small and statistically insignificant.

Based on our theoretical framework, we would expect received transfers to also be affected by the NHIS implementation. One economic explanation for our empirical findings could be asymmetric information between network partners. As the NHIS was disseminated gradually, individuals who were already covered by the insurance might still have received transfers from districts where the NHIS was not available.

Table 5.5 Effect of the NHIS implementation on receiving transfers

	LPM 1	LPM 2	LPM 3	OLS 1	OLS 2	OLS 3
NHIS	-0.089	-0.081	-0.086	-1.926	-0.835	-0.640
	(0.092)	(0.106)	(0.105)	(14.649)	(14.919)	(15.247)
N	4985	4985	4985	4985	4985	4985
adj. R-sq	0.040	0.091	0.091	0.007	0.065	0.065
District and interview month dummies	Yes	Yes	Yes	Yes	Yes	Yes
Individual and HH control variables	No	Yes	Yes	No	Yes	Yes
Urban time trend	No	No	Yes	No	No	Yes

Notes: Standard errors (in parentheses) are clustered at the district level. * p <0.10, ** p <0.05, *** p <0.01; See Appendix 5.A4 for detailed regression outputs.

Unfortunately, we cannot test this hypothesis, as we cannot determine which districts the donor and recipient from the same network live in; however, a growing body of empirical literature has shown that information asymmetries play a crucial role in remittance decisions (Ashraf, 2009; Jakiela and Ozier, 2012; Ambler, 2015).

We next assess whether the size of the crowding out depends on the relationship status and sharing obligations between network partners. In order to examine whether this factor mediates the impact of the NHIS implementation on informal transfer behaviour, the next subsection explores the heterogeneity of the NHIS effect using information on the relationship status between donor and recipient.

5.4.1 Treatment Heterogeneity

In order to empirically examine the role of kinship relationships in our empirical findings, we explore the heterogeneity of the NHIS effect using information on the relationship status between donor and recipient. We define three relationship groups: children/parents (inner family), siblings/cousins/uncles and aunts/grandparents (extended family) and non-relatives.

First, we use the information on whether a respondent received transfers from these three groups as dependent variables and estimate the group-specific effects of implementing the NHIS (see Table 5.6). The probability of receiving transfers is negative and statistically significant if the recipient receives transfers from non-relatives. Furthermore, the estimates show that the probability of receiving less is reduced the closer the relationship between recipient and donor, which is in line with the second prediction of our theoretical framework: that the degree of the crowding out depends on the relationship status and sharing obligations (exit costs) between network participants. Thus, it seems that the closer the relationship (strong sharing

obligations), the lower the crowding out.

Table 5.6 Effect of the NHIS implementation on received transfers from different donors

Transfers received from:	Children/ Parents LPM 1	Ext. Family/ Siblings LPM 2	Non-relatives LPM 3	Children/ Parents OLS 1	Ext. Family/ Siblings OLS 2	Non-relatives OLS 3
NHIS	-0.071	-0.141	-0.155*	13.884	-4.847	-14.625
	(0.111)	(0.112)	(0.088)	(15.794)	(16.289)	(9.062)
N	4018	3590	3706	4018	3590	3706
adj. R-sq	0.158	0.071	0.078	0.041	0.044	0.037

Notes: Individual and household control variables are included. District and month dummies are included. Standard errors (in parentheses) are clustered at the district level. * p <0.10, ** p <0.05, *** p <0.01

The results at the intensive margin for receiving transfers from different donors show a positive but statistically insignificant effect for children/parents as donors. Conversely, we find a negative effect for non-relatives which is close to being statistically significant at the 10 % level. As the share of received transfers from inner family members (children/parents) is similar to the share of received transfers from non-relatives (see Appendix 5.A1), the effects offset each other on average. Thus, in combination with the asymmetric information between network partners, heterogeneity in the treatment effect (i.e., the availability of the NHIS) can explain why we only found a small effect of NHIS implementation on transfers received.

Second, we use the information on whether a respondent sent transfers to children/parents (inner family), extended family members and non-relatives as dependent variables (see Table 5.7). Compared to transfers received, we find the same kind of pattern for the NHIS effect across the three groups at the extensive and intensive margins. Interestingly, we do not find the same offsetting effect as with received transfers, as donors substantially reduced remittances to non-relatives after implementation of the NHIS. Thus, the crowding out effect is mainly driven by a short-term reduction in remittances to non-relatives. In addition, the coefficients for sent transfers are lower than those for received transfers, which again emphasizes the importance of asymmetric information on insurance status across districts. If this is the case, recipients who were already covered by the NHIS would still receive transfers from network partners who were not aware that the NHIS had been implemented in their network partner's district.

As the NHIS covers all basic outpatient/inpatient services and also treatment for chronic diseases, it is likely that the crowding out effects are largest for individuals

Table 5.7 Effect of the NHIS implementation on transfers sent to different recipients

Transfers sent to:	Children/ Parents LPM 1	Ext. Family/ Siblings LPM 2	Non-relatives LPM 3	Children/ Parents OLS 1	Ext. Family/ Siblings OLS 2	Non-relatives OLS 3
NHIS	-0.081	-0.182**	-0.259***	1.821	-6.158	-38.331*
	(0.085)	(0.086)	(0.078)	(19.562)	(10.102)	(22.619)
N	4136	3623	3813	4136	3623	3813
adj. R-sq	0.114	0.154	0.186	0.144	0.095	0.051

Notes: Individual and household control variables are included. District and month dummies are included. Standard errors (in parentheses) are clustered at the district level. * p <0.10, ** p <0.05, *** p <0.01

sending/receiving regular transfers for health purposes. To investigate whether this is indeed the case, we use information on the purposes of transfers. 58 percent of respondents indicated 'health' as one of the main two reasons for sending/receiving regular transfers. Table 5.8 contains estimated coefficients for a regression model using the values of regular transfers made for different purposes as dependent variables. The results indicate that, after the implementation of the NHIS, donors mainly reduced transfers made for health reasons, which is in line with the negative impact of the NHIS on OOP expenditures (see Table 5.3).

Table 5.8 Effect of the NHIS implementation on transfers sent for different purposes

Dep. var.:	Health	Daily consumption	Education	Housing
NHIS	-28.785**	-1.048	-12.413	-4.496
	(14.512)	(5.317)	(14.028)	(5.694)
N	3849	3422	3784	3278
adj. R-sq	0.101	0.060	0.142	0.062

Notes: Individual and household control variables are included. District and month dummies are included. Standard errors (in parentheses) are clustered at the district level. * p <0.10, ** p <0.05, *** p <0.01

5.4.2 Robustness Checks

We next examine whether the sequencing of the NHIS implementation within districts, i.e., on the sub-district level, confounds our identification strategy. Following our baseline specification (equation (5.1)), we conduct regressions using time-

invariant and predetermined characteristics such as education level, gender and regional characteristics as dependent variables. The results show that the implementation of the NHIS has no effect on these time-invariant and pre-determined characteristics (see 5.A5 and 5.A6 in the Appendix). We also estimate the effect of the NHIS on our main outcome variables by using the fourth wave of the GLSS to test whether our findings are confounded by the timing of interviews between sub-districts. As this wave was conducted in the same manner and contains the same districts, we can adapt the NHIS variable for that time and provide a placebo estimate. The results show that the implementation of the NHIS had no significant effect on the probability of sending or receiving transfers (see Appendix 5.A7). In addition, no significant effect was found for the monetary equivalents.

Overall, we conclude that our findings are due to the implementation of the NHIS and are not driven by a systematic relationship between the NHIS rollout, sub-districts, interview dates and the outcome variables.

As a certain amount of insurance premiums had to be collected before the NHIS was officially launched in each district, most individuals had to pay premiums before actually being able to benefit from insurance coverage. Thus, we explore whether individuals changed their transfer behaviour in anticipation of the official launch of the NHIS. If this was the case, our estimates could be biased. To investigate the presence of this bias, we 'shift' our treatment indicator backwards by two months. The results indicate no significant changes in transfer behaviour before the scheme was officially launched (see 5.A8 in the Appendix of this chapter). Thus, changes in transfer behaviour are only found once district authorities had officially launched the NHIS and healthcare services were freely available for premium payers. This suggests that the respondents did not substitute premium payments by reducing remittance to other households.

5.5 Conclusion

This chapter provided empirical evidence that a formal health insurance scheme crowds out regular informal transfers in Ghana. We analysed cross-sectional data from the fifth Ghanaian Living Standard Survey, which benefited from the fact that various sub-districts were surveyed at different times during the survey period. As most districts introduced the NHIS during this period, we compared different individuals who were interviewed before and after the introduction of the insurance scheme. We evaluated whether the availability of formal insurance resulted in re-

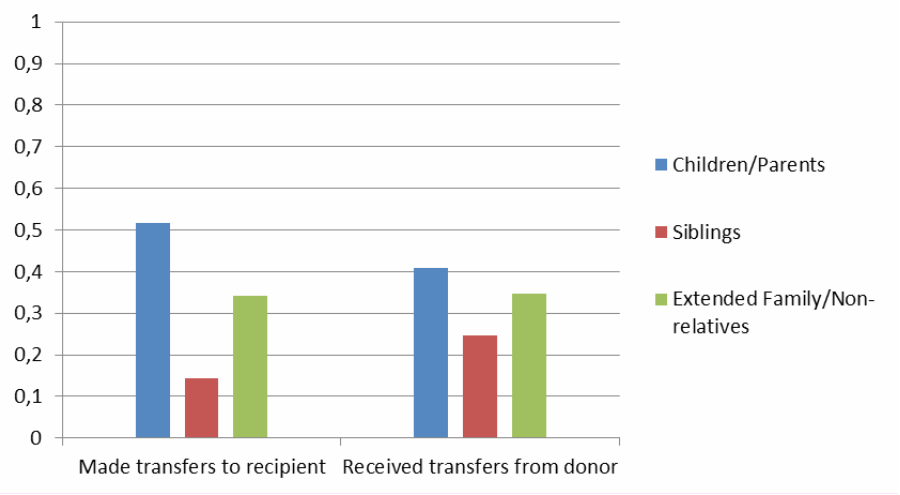
duced contributions to informal transfer networks and also investigated the impact of the NHIS on health-related outcomes. We also investigated whether our results were due to the relationship status and level of sharing obligations between network partners. We differentiated between three groups: Children/parents (inner family), siblings/cousins/aunts and uncles (extended family) and non-relatives. We also explored whether the results were driven by different reasons for making transfers. Our findings suggest that there is a crowding out effect, since the introduction of the formal health insurance scheme reduced the probability of making transfers. The value of remittances also decreased to a significant extent, due to a reduction in transfers for health purposes. Our analysis of health-related outcomes found that the NHIS reduced respondents' OOP expenditures. Interestingly, the reduction in expenditures was very similar to the reduction in sent transfers. We also found that the closer the relationship between recipient and donor, the lower the crowding out. Thus, the degree of crowding out depends strongly on the relationship status and sharing obligations between network partners. As the NHIS covers all basic outpatient and inpatient services, such as blood tests and malaria treatments, and also maternity care services, we interpret our results as indicating that it is not only ill individuals who benefit financially from the NHIS: donors have also been financially relieved by the implementation of the NHIS.

This effect seems most relevant for unrelated network partners, as they showed the largest reduction of transfers. Lower sharing obligations and information asymmetries make it possibly less costly for these network members to reduce transfer payments after the implementation of the NHIS. Since the risk of default in times of financial hardship due to health shocks is probably higher in networks characterized by low sharing obligations, it is more beneficial for these individuals to reduce transfers and rely on formal insurance mechanisms. Conversely, as we do not find large crowding out effects for networks between immediate family members, we conclude that it probably takes more time to convince members of networks characterized by strong sharing obligations and high costs of leaving the network.

As we only observe changes in transfer behaviour in the short run, we believe the effects of changes in investments or savings are likely to take more time to become apparent. However, from a policy perspective it would be useful to investigate whether the observed changes in transfer behaviour translate into higher investments or savings in the long run. In particular, it would be interesting to investigate the extent to which the crowding out of informal transfers is used for investments or consumption purposes by also considering the direct (insurance premiums) and indirect

costs (2.5% addition to the value added tax) of the NHIS. As the sixth round of the GLSS is now available, a promising avenue for future research would be to examine whether the implementation of the NHIS has on average been a net gain or loss for covered individuals in the long run.

Appendix Chapter 5



5.A1 Shares of sent and received transfers from/for different donors/recipients

5.A2 Descriptive statistics

Variable	Sent transfers Mean	Received transfers Mean
Household size	5.29	5.26
HH expenditures quintile 2 (0/1)	0.2	0.2
HH expenditures quintile 3 (0/1)	0.19	0.18
HH expenditures quintile 4 (0/1)	0.19	0.17
HH expenditures quintile 5 (0/1)	0.22	0.18
HH savings account (0/1)	0.29	0.23
Migrant (0/1)	0.18	0.18
Formal employment (0/1)	0.13	0.09
Informal employment (0/1)	0.06	0.06
Self employment (0/1)	0.79	0.81
Primary school (0/1)	0.16	0.16
Junior high school (0/1)	0.17	0.17
Secondary high school (0/1)	0.27	0.23
Technical school (0/1)	0.05	0.04
University (0/1)	0.02	0.01
Female (0/1)	0.54	0.55
Age	37.04	0.61
Married (0/1)	0.61	37.48
Urban (0/1)	0.32	0.28
Number of observations	5956	4985

5.A3 Effect on sent transfers (detailed output)

	LPM 1	LPM 2	LPM 3	OLS 1	OLS 2	OLS 3
NHIS	-0.148*	-0.146**	-0.148**	-23.803	-24.172*	-23.723*
	(0.079)	(0.061)	(0.060)	(14.866)	(14.130)	(14.005)
Household size		0.018***	0.018***		5.395***	5.390***
		(0.004)	(0.004)		(1.197)	(1.196)
HH expenditures Q2		0.086***	0.085***		26.224***	26.318***
		(0.031)	(0.031)		(7.742)	(7.741)
HH expenditures Q3		0.117***	0.116***		40.484***	40.671***
		(0.037)	(0.037)		(10.417)	(10.390)
HH expenditures Q4		0.134***	0.133***		52.712***	52.933***
		(0.039)	(0.039)		(9.503)	(9.472)
HH expenditures Q5		0.263***	0.264***		111.333***	111.184***
		(0.041)	(0.041)		(11.463)	(11.395)
HH savings account		0.138***	0.139***		41.733***	41.469***
		(0.022)	(0.023)		(6.962)	(6.989)
Migrant		-0.004	-0.002		10.388	10.040
		(0.023)	(0.023)		(8.321)	(8.288)
Formal employment		0.343***	0.347***		69.156***	68.382***
		(0.041)	(0.041)		(13.421)	(13.479)
Informal employment		0.099**	0.099**		8.514	8.464
		(0.044)	(0.044)		(7.622)	(7.663)
Self employment		0.211***	0.214***		56.270***	55.815***
		(0.029)	(0.028)		(9.284)	(9.252)
Primary school		-0.007	-0.007		0.495	0.462
		(0.023)	(0.023)		(5.602)	(5.617)
Junior high school		0.002	0.000		5.543	5.846
		(0.019)	(0.019)		(5.565)	(5.689)
Secondary high school		0.030*	0.029		19.585***	19.798***
		(0.018)	(0.018)		(5.408)	(5.426)
Technical school		0.046	0.046		38.429***	38.441***
		(0.032)	(0.032)		(11.703)	(11.702)
University		0.067	0.064		95.394***	95.807***
		(0.042)	(0.043)		(22.007)	(22.052)
Female		-0.036***	-0.036***		-4.623	-4.543
		(0.010)	(0.010)		(2.811)	(2.819)
Married		0.042***	0.041***		2.843	3.009
		(0.014)	(0.014)		(4.224)	(4.226)
Age		0.005*	0.005*		0.998	0.992
		(0.003)	(0.003)		(0.817)	(0.816)
Age squared		-0.000**	-0.000**		-0.013	-0.012
		(0.000)	(0.000)		(0.009)	(0.009)
Urban		-0.117***	-0.165***		-14.373	-5.313
		(0.032)	(0.057)		(10.588)	(17.425)
Urban time trend			0.009			-1.700
			(0.008)			(2.443)
N	5956	5956	5956	5956	5956	5956
adj. R-sq	0.061	0.163	0.164	0.023	0.150	0.150

Notes: Individual and household control variables are included. District and month dummies are included. Standard errors (in parentheses) are clustered at the district level. * p <0.10, ** p <0.05, *** p <0.01

5.A4 Effect on received transfers (detailed output)

	LPM 1	LPM 2	LPM 3	OLS 1	OLS 2	OLS 3
NHIS	-0.089	-0.081	-0.086	-1.926	-0.835	-0.640
	(0.092)	(0.106)	(0.105)	(14.649)	(14.919)	(15.247)
Household size		0.001	0.001		1.299	1.294
		(0.004)	(0.004)		(0.992)	(0.996)
HH expenditures Q2		0.008	0.008		7.634	7.651
		(0.036)	(0.036)		(5.772)	(5.783)
HH expenditures Q3		-0.011	-0.011		12.063	12.084
		(0.032)	(0.032)		(8.118)	(8.139)
HH expenditures Q4		0.038	0.038		28.051***	28.060***
		(0.035)	(0.036)		(10.149)	(10.139)
HH expenditures Q5		0.062	0.065		37.260***	37.169***
		(0.044)	(0.044)		(11.628)	(11.641)
HH savings account		0.002	0.001		-1.829	-1.818
		(0.025)	(0.025)		(5.365)	(5.351)
Migrant		-0.050**	-0.049**		-10.681*	-10.705*
		(0.023)	(0.023)		(5.398)	(5.397)
Formal employment		-0.355***	-0.350***		-112.921***	-113.099***
		(0.052)	(0.052)		(22.358)	(21.980)
Informal employment		-0.287***	-0.286***		-98.208***	-98.264***
		(0.066)	(0.065)		(27.237)	(27.061)
Self employment		-0.303***	-0.301***		-110.627***	-110.687***
		(0.054)	(0.053)		(24.165)	(24.001)
Primary school		-0.001	-0.001		2.989	2.995
		(0.019)	(0.019)		(4.014)	(4.027)
Junior high school		0.033	0.031		8.671	8.738
		(0.022)	(0.023)		(5.512)	(5.576)
Secondary high school		0.010	0.007		13.813**	13.905**
		(0.025)	(0.025)		(6.213)	(6.340)
Technical school		0.043	0.042		28.391**	28.432**
		(0.041)	(0.041)		(12.034)	(12.125)
University		-0.019	-0.020		11.981	12.022
		(0.045)	(0.046)		(15.392)	(15.398)
Female		0.065***	0.065***		22.136***	22.137***
		(0.013)	(0.013)		(4.115)	(4.116)
Married		-0.059***	-0.060***		-7.755*	-7.724*
		(0.017)	(0.016)		(4.597)	(4.548)
Age		-0.011***	-0.011***		-2.402***	-2.405***
		(0.003)	(0.003)		(0.622)	(0.626)
Age squared		0.000***	0.000***		0.032***	0.032***
		(0.000)	(0.000)		(0.008)	(0.008)
Urban		-0.020	-0.074		-7.809	-5.806
		(0.033)	(0.067)		(6.941)	(16.660)
Urban time trend			0.010			-0.360
			(0.011)			(2.645)
N	4985	4985	4985	4985	4985	4985
adj. R-sq	0.043	0.091	0.091	0.007	0.065	0.065

Notes: Individual and household control variables are included. District and month dummies are included. Standard errors (in parentheses) are clustered at the district level. * p <0.10, ** p <0.05, *** p <0.01

5.A5 Effect of the NHIS implementation on time-invariant characteristics 1

	Urban	Female	Married	Primary School	Jun. High School	Secondary School	Technical School	University
NHIS	-0.039	0.019	-0.184	-0.027	0.037	-0.024	-0.017	-0.017
	(0.122)	(0.038)	(0.177)	(0.033)	(0.029)	(0.035)	(0.022)	(0.019)
N	5956	5956	5956	5956	5956	5956	5956	5956
adj. R-sq	0.112	0.002	0.002	0.002	0.002	0.006	0.013	0.004

Notes: Sample of made transfers is used. District and month dummies are included. Standard errors (in parentheses) are clustered at the district level. * p <0.10, ** p <0.05, *** p <0.01

5.A6 Effect of the NHIS implementation on time-invariant characteristics 2

	Urban	Female	Married	Primary School	Jun. High School	Secondary School	Technical School	University
NHIS	-0.139	0.036	-0.027	0.006	0.025	-0.021	0.011	-0.005
	(0.118)	(0.035)	(0.056)	(0.034)	(0.041)	(0.051)	(0.020)	(0.007)
N	4985	4985	4985	4985	4985	4985	4985	4985
adj. R-sq	0.060	0.001	0.001	0.002	0.001	0.006	0.011	0.001

Notes: Sample of received transfers is used. District and month dummies are included. Standard errors (in parentheses) are clustered at the district level. * p <0.10, ** p <0.05, *** p <0.01

5.A7 Placebo effect of the NHIS implementation

Variables	Transfers made (1/0)	Amount of made transf.	Received transfers (1/0)	Amount of received transf.	Low health status (1/0)	# of sick days	OOP expend.
NHIS	0.19	6.41	-0.05	-6.33	-0.04	-0.29	0.19
	(0.15)	(4.57)	(0.04)	(19.80)	(0.05)	(0.42)	(0.26)
N	5741	5741	5929	5929	5741	5741	5741

Notes: GLSS 1998/1999 is used. Individual and household variables are included. District and month dummies are included. Standard errors (in parenthesis) are clustered at the district level. * p <0.10, ** p <0.05, *** p <0.01

5.A8 NHIS shift

	Sent Transfers		Received Transfers	
	LPM (t-2)	LPM (t-1)	LPM (t-2)	LPM (t-1)
NHIS	-0.033	-0.10	0.022	-0.085
	(0.095)	(0.101)	(0.122)	(0.146)
Household size	0.082***	0.080**	-0.002	-0.005
	(0.031)	(0.031)	(0.035)	(0.036)
HH expenditures Q2	0.116***	0.115***	-0.014	-0.016
	(0.036)	(0.036)	(0.033)	(0.034)
HH expenditures Q3	0.132***	0.130***	0.023	0.021
	(0.038)	(0.039)	(0.037)	(0.037)
HH expenditures Q4	0.257***	0.256***	0.075*	0.073
	(0.040)	(0.040)	(0.045)	(0.045)
HH expenditures Q5	0.042***	0.043***	-0.057***	-0.057***
	(0.014)	(0.014)	(0.016)	(0.017)
HH savings account	0.135***	0.135***	0.021	0.021
	(0.022)	(0.022)	(0.025)	(0.026)
Migrant	0.018***	0.018***	0.003	0.002
	(0.004)	(0.004)	(0.004)	(0.004)
Formal employment	0.034	0.034	0.021	0.019
	(0.021)	(0.021)	(0.021)	(0.021)
Informal employment	0.003	0.001	-0.047**	-0.049**
	(0.022)	(0.022)	(0.022)	(0.022)
Self employment	0.346***	0.349***	-0.334***	-0.329***
	(0.041)	(0.041)	(0.051)	(0.051)
Primary school	0.099**	0.102**	-0.299***	-0.296***
	(0.043)	(0.043)	(0.057)	(0.057)
Jun. high school	0.210***	0.214***	-0.291***	-0.287***
	(0.026)	(0.027)	(0.048)	(0.049)
Sec. high school	-0.007	-0.008	0.002	0.002
	(0.023)	(0.023)	(0.020)	(0.020)
Technical school	-0.002	-0.001	0.031	0.031
	(0.019)	(0.019)	(0.023)	(0.023)
University	0.029	0.028	0.021	0.022
	(0.017)	(0.017)	(0.026)	(0.026)
Female	0.047	0.046	0.061	0.063
	(0.031)	(0.031)	(0.044)	(0.044)
Married	0.036***	0.036***	-0.067***	-0.068***
	(0.009)	(0.009)	(0.013)	(0.013)
Age	0.004	0.004	-0.012***	-0.012***
	(0.003)	(0.003)	(0.003)	(0.003)
Age squared	-0.000**	-0.000**	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
N	5956	5956	4985	4985

Notes: District and month dummies are included. Standard errors (in parentheses) are clustered at the district level. * p <0.10, ** p <0.05, *** p <0.01

Chapter 6

Closing Remarks

This dissertation assessed contemporary health economics research questions with an emphasis on healthcare expenditures. We emphasised the importance of ageing in properly predicting healthcare expenditures and the role of individual characteristics that may be barriers to successful participation in public healthcare programmes. We also critically discussed the methods used to detect selection and information asymmetries in insurance markets and evaluated the impact of a formal insurance scheme on health and participation in informal transfer networks.

First, the role of ageing and end-of-life morbidity for projecting healthcare expenditures was evaluated using aggregated data from Sweden. We provided an indicator capturing severe morbidity at the end of life to empirically distinguish a pure ageing effect from morbidity in the context of (non-curative) long-term care expenditures. This empirical application is of special interest in the context of countries where costs of ageing are expected to be considerable due to an increasing share of the elderly in the population.

Second, using survey data, we evaluated whether restrictions in individual health decision-making processes can serve as barriers to participating in a national healthcare programme in England. If the success of a healthcare policy depends on the behaviour of individual decision makers, policy designers must carefully account for these individual characteristics to achieve the desired programme outcomes. We empirically showed that a decline in an individual's ability to memorise is negatively related to their decision to participate in a public healthcare programme against bowel cancer. We interpret this finding as an indication that participation with the programme is suboptimal from a normative point of view if we assume that the people in our analysis would have participated if their memory had not declined. From an applied perspective, our findings are of interest because health policy makers should

be able to estimate participation in a healthcare programme to be able to judge the corresponding consequences for (public) healthcare expenditures.

Third, the dissertation contributed to detecting selection in health insurance markets. Efficiency is a major concern in the stability of private health insurance markets across both developing an developed countries, since (additional) private health insurance is potentially an important part of the architecture of future healthcare systems. However, if selection occurs in health insurance markets based on risk which is not accounted for in risk premiums, this suggests inefficient resource allocation, potentially inducing market instability. We provided further knowledge about commonly applied approaches used to detect selection in insurance markets based on 'unused characteristics'. We show that these approaches can easily lead to false conclusions about the direction of selection, i.e., whether adverse or advantageous selection is occurring. This knowledge should be used to empirically assess selection in insurance markets and derive consequences for efficiency and market stability.

Fourth, we assessed the relationship between public health insurance and private transfer networks from an economic development perspective. Using survey data from Ghana, we analysed whether the introduction of a nationwide formal health insurance scheme reduced participation in informal transfers. We found robust evidence that the availability of the insurance scheme reduced transfers at the extensive and intensive margin. This finding may be beneficial from a development perspective, since the crowding out may reduce the high financial burden for employers. However, our estimates suggest that this crowding out depends strongly on the relationship status between network members. Since the crowding out of informal transfers at the intensive margin is highest for transfers to non-relatives, a group where sharing obligations are probably relatively low, overoptimistic expectations about positive effects due to a reduction in social pressure within transfer networks should be mitigated. In addition, we found insurance reduced OOP health payments. Under the assumption that resource allocation for health purposes is organised more efficiently under a formal insurance scheme, our findings indicate that formal health insurance can be beneficial from a development perspective.

One lesson that can be learned from this dissertation is that when mean outcomes are emphasised in an empirical analysis, taking a closer look at subgroups can often be fruitful for assessing heterogeneous individual behaviour or heterogeneous policy effects. This knowledge can be used to provide a better understanding of human behaviour and successful policy design. For example, when estimating the role of ageing and morbidity on LTC expenditures, we found heterogeneity for age/sex outcomes

in different domains of LTC provision. We also found evidence that heterogeneous effects play a role when analysing the relationship between formal insurance and informal transfer networks. In the context of detecting selection in insurance markets, it is clear that conclusions drawn using the discussed methods can be completely misleading if group specific heterogeneity is not accounted for.

Our findings may be further used to test and develop hypotheses to generate knowledge about decision making in many health economics contexts and to design successful policies. Proper policy design and evaluation are key, because increasing healthcare expenditures and financial restrictions in public budgets may induce substantial burden for financing welfare states.

When detecting selection in insurance markets, future research may allow for group-specific parameter heterogeneity to decrease the data dimension, which may be advantageous from an empirical perspective and provide further insights into selection mechanisms. We also see the relationship between formal and informal markets in the context of healthcare provision and insurance as an important research question in more developed countries. We hope our ideas and contributions will be helpful for future research and promote the successful adjustment and design of healthcare policies, allowing for both successful and affordable healthcare provision in different regions of the world.

Bibliography

Akerlof, G. (1970). The market for 'lemons': quality uncertainty and the market mechanism. *The Quarterly Journal of Economics 84*(3), 488–500.

Ambler, K. (2015). Don't tell on me: experimental evidence of asymmetric information in transnational households. *Journal of Development Economics 113*, 52–69.

Angrist, J. D. and J.-S. Pischke (2009). *Mostly harmless econometrics: an empiricist's companion*, Volume 1. Princeton University Press.

Arentz, C., J. Eekhoff, and S. Kochskämper (2012). Private health insurance: a role model for European health systems. *The European Journal of Health Economics 13*(5), 615–621.

Ashraf, N. (2009). Spousal control and intra-household decision making: an experimental study in the Philippines. *The American Economic Review 99*(4), 1245–1277.

Attanasio, O. and J.-V. Rıos-Rull (2000). Consumption smoothing in island economies: Can public insurance reduce welfare? *European Economic Review 44*(7), 1225–1258.

Avitabile, C., T. Jappelli, and M. Padula (2011). Cognitive abilities, healthcare and screening tests. *Journal of Population Ageing 4*(4), 251–269.

Bauhoff, S. (2012). Do health plans risk-select? An audit study on Germany's social health insurance. *Journal of Public Economics 96*(9), 750–759.

Bech, M., T. Christiansen, E. Khoman, J. Lauridsen, and M. Weale (2011). Ageing and health care expenditure in EU-15. *The European Journal of Health Economics 12*(5), 469–478.

Bolhaar, J., M. Lindeboom, and B. van der Klaauw (2012). A dynamic analysis of the demand for health insurance and health care. *European Economic Review 56*(4), 669–690.

Bolin, K., D. Hedblom, A. Lindgren, and B. Lindgren (2010). Asymmetric information and the demand for voluntary health insurance in Europe. Technical report, National Bureau of Economic Research.

Boyle, S. (2011). United Kingdom (England) health system review. *Health systems in transition 13*(1).

Breyer, F., N. Lorenz, and T. Niebel (2015). Health care expenditures and longevity: Is there a Eubie Blake effect? *The European Journal of Health Economics 16*(1), 95–112.

Brune, L., X. Giné, J. Goldberg, and D. Yang (2015). Facilitating savings for agriculture: field experimental evidence from Malawi. Technical report, National Bureau of Economic Research.

Buchmueller, T. C., D. G. Fiebig, G. Jones, and E. Savage (2013). Preference heterogeneity and selection in private health insurance: the case of Australia. *Journal of Health Economics 32*(5), 757–767.

Centralbyrån, S.-S. (2009). Tabeller över sveriges befolkning 2006.

Chandra, A. and J. Skinner (2012). Technology growth and expenditure growth in health care. *Journal of Economic Literature 50*(3), 645–80.

Chetty, R. (2015). Behavioral economics and public policy: a pragmatic perspective. *American Economic Review 105*(5), 1–33.

Chiappori, P. and B. Salanié (1997). Empirical contract theory: the case of insurance data. *European Economic Review 41*(3-5), 943–950.

Cohen, A. and P. Siegelman (2010). Testing for adverse selection in insurance markets. *Journal of Risk and Insurance 77*(1), 39–84.

Colombo, F., L. Ana, M. Jérôme, and T. Frits (2011). *OECD Health Policy Studies Help Wanted?: Providing and Paying for Long-Term Care*. Organisation for Economic Co-Operation and Development.

Colombo, F. and N. Tapay (2004). Private health insurance in OECD countries. *OECD Health Working Papers, No. 15*.

Cox, D. (1987). Motives for private income transfers. *The Journal of Political Economy 95*(3), 508–546.

Crump, R. K., V. J. Hotz, G. W. Imbens, and O. A. Mitnik (2009). Dealing with limited overlap in estimation of average treatment effects. *Biometrika*, 1–13.

Culyer, A. J. and A. Wagstaff (1993). Equity and equality in health and health care. *Journal of Health Economics 12*(4), 431–457.

Cutler, D. M., A. Finkelstein, and K. McGarry (2008). Preference heterogeneity and insurance markets: explaining a puzzle of insurance. *The American Economic Review 98*(2), 157–62.

Cutler, D. M. and E. Glaeser (2005). What explains differences in smoking, drinking, and other health-related behaviors? *The American Economic Review 95*(2), 238–242.

Cutler, D. M. and A. Lleras-Muney (2010). Understanding differences in health behaviors by education. *Journal of Health Economics 29*(1), 1–28.

Cutler, D. M. and R. J. Zeckhauser (2000). The anatomy of health insurance. *Handbook of Health Economics 1*, 563–643.

Cylus, J., E. Richardson, L. Findley, M. Longley, C. O'Neill, and D. Steel (2015). United Kingdom: Health system review. *Health systems in transition 17*(5), 1–126.

de Meijer, C., M. Koopmanschap, T. B. d'Uva, and E. Van Doorslaer (2011). Determinants of long-term care spending: age, time to death or disability? *Journal of Health Economics 30*(2), 425–438.

De Weerdt, J. and M. Fafchamps (2011). Social identity and the formation of health insurance networks. *Journal of Development Studies 47*(8), 1152–1177.

Deaton, A. (1997). *The analysis of household surveys: a microeconometric approach to development policy*. World Bank Publications.

Decker, S. and H. Schmitz (2015). Health shocks and risk aversion. *Ruhr Economic Paper* (581).

Dercon, S. and P. Krishnan (2003). Risk sharing and public transfers. *The Economic Journal 113*(486), C86–C94.

Di Falco, S. and E. Bulte (2011). A dark side of social capital? Kinship, consumption, and savings. *Journal of Development Studies 47*(8), 1128–1151.

Doiron, D., G. Jones, and E. Savage (2008). Healthy, wealthy and insured? The role of self-assessed health in the demand for private health insurance. *Health Economics 17*(3), 317–334.

Dubois, P., B. Jullien, and T. Magnac (2008). Formal and informal risk sharing in LDCs: theory and empirical evidence. *Econometrica 76*(4), 679–725.

Duflo, E., M. Kremer, and J. Robinson (2009). Nudging farmers to use fertilizer: theory and experimental evidence from Kenya. Technical report, National Bureau of Economic Research.

Dye, C., J. C. Reeder, and R. F. Terry (2013). *Research for universal health coverage*. World Health Organization.

Eide, E. R. and M. H. Showalter (2011). Estimating the relation between health and education: What do we know and what do we need to know? *Economics of Education Review 30*(5), 778–791.

Einav, L. and A. Finkelstein (2011). Selection in insurance markets: theory and empirics in pictures. *The Journal of Economic Perspectives 25*(1), 115–138.

Einav, L., A. Finkelstein, and J. Levin (2009). Beyond testing: empirical models of insurance markets. *Annual Review of Economics 2*, 311–336.

Esping-Andersen, G. (1990). *The Three Worlds of Welfare Capitalism*. Princeton University Press.

Fafchamps, M. (2008). Risk sharing between households. *Handbook of Social Economics 1*.

Fang, H., M. Keane, A. Khwaja, M. Salm, and D. Silverman (2007). Testing the mechanisms of structural models: the case of the Mickey Mantle effect. *The American Economic Review 97*(2), 53–59.

Fang, H., M. P. Keane, and D. Silverman (2008). Sources of advantageous selection: evidence from the medigap insurance market. *Journal of Political Economy 116*(2), 303–350.

Farag, M., A. NandaKumar, S. Wallack, D. Hodgkin, G. Gaumer, and C. Erbil (2012). The income elasticity of health care spending in developing and developed countries. *International Journal of Health Care Finance and Economics 12*(2), 145–162.

Felder, S., A. Werblow, and P. Zweifel (2010). Do red herrings swim in circles? Controlling for the endogeneity of time to death. *Journal of Health Economics 29*(2), 205–212.

Finkelstein, A. and K. McGarry (2006). Multiple dimensions of private information: evidence from the long-term care insurance market. *American Economic Review 96*(4), 938–958.

Finkelstein, A. and J. Poterba (2014). Testing for asymmetric information using unused observables in insurance markets: evidence from the UK annuity market. *Journal of Risk and Insurance 81*(4), 709–734.

Forma, L., P. Rissanen, A. Noro, J. Raitanen, and M. Jylhä (2007). Health and social service use among old people in the last 2 years of life. *European Journal of Ageing 4*(3), 145–154.

Frey, B. (1990). *Ökonomie ist Sozialwissenschaft: die Anwendung der Ökonomie auf neue Gebiete*. F. Vahlen.

Fries, J. F. (1980). Aging, natural death, and the compression of morbidity. *New England Journal of Medicine 303*(3), 130–135.

Fukushima, N., J. Adami, and M. Palme (2010). The long-term care system for the elderly in Sweden. Technical report, ENEPRI Research Report.

García-Gómez, P., A. M. Jones, and N. Rice (2010). Health effects on labour market exits and entries. *Labour Economics 17*(1), 62–76.

Getzen, T. (2000a). Forecasting health expenditures: short, medium and long (long) term. *Journal of Health Care Finance 26*(3), 56–72.

Getzen, T. (2000b). Health care is an individual necessity and a national luxury: applying multilevel decision models to the analysis of health care expenditures. *Journal of Health Economics 19*(2), 259–270.

Getzen, T. (2006). Aggregation and the measurement of health care costs. *Health Services Research 41*(5), 1938–1954.

Glied, S. and A. Lleras-Muney (2008). Technological innovation and inequality in health. *Demography 45*(3), 741–761.

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology 78*(6), 1360–1380.

Gremeaux, V., M. Gayda, R. Lepers, P. Sosner, M. Juneau, and A. Nigam (2012). Exercise and longevity. *Maturitas 73*(4), 312–317.

Grimm, M., R. Hartwig, and J. Lay (2013). Does forced solidarity hamper investment in small and micro enterprises? *IZA Discussion Paper No. 7229*.

Hayes, T. L., N. Larimer, A. Adami, and J. A. Kaye (2009). Medication adherence in healthy elders: small cognitive changes make a big difference. *Journal of Ageing and Health 21*(4), 567–580.

Herzer, D., H. Strulik, and S. Vollmer (2012). The long-run determinants of fertility: one century of demographic change 1900–1999. *Journal of Economic Growth 17*(4), 357–385.

Hoff, K. and A. Sen (2005). The kin system as a poverty trap? *World Bank Policy Research Working Paper* (3575).

Idler, E. and Y. Benyamini (1997). Self-rated health and mortality: a review of twenty-seven community studies. *Journal of Health and Social Behavior 38*(1), 21–37.

Insel, K. C., G. O. Einstein, D. G. Morrow, and J. T. Hepworth (2013). A multifaceted prospective memory intervention to improve medication adherence: design of a randomized control trial. *Contemporary Clinical Trials 34*(1), 45–52.

Jakiela, P. and O. W. Ozier (2012). Does Africa need a rotten kin theorem? Experimental evidence from village economies. *World Bank Policy Research Working Paper* (6085).

Johar, M., G. Jones, M. Keane, E. Savage, and O. Stavrunova (2013). The demand for private health insurance: Do waiting lists matter?-revisited. *Working Paper*.

Kabki, M. (2007). *Transnationalism, local development and social security: the functioning of support networks in rural Ghana*. African Studies Centre, Leiden.

Kaestner, R. (2009). Adolescent cognitive and non-cognitive correlates of adult health. Technical report, National Bureau of Economic Research.

Kahneman, D. (2003). Maps of bounded rationality: psychology for behavioral economics. *American Economic Review 93*(5), 1449–1475.

Karlsson, M., T. Iversen, and H. Øien (2010). Scandinavian long-term care financing. Technical report, Oslo University, Health Economics Research Programme.

Karlsson, M. and F. Klohn (2011). Ageing, health and disability – An economic perspective. In *From Grey to Silver*, pp. 51–67. Springer.

Karlsson, M. and F. Klohn (2014). Testing the red herring hypothesis on an aggregated level: ageing, time-to-death and care costs for older people in Sweden. *The European Journal of Health Economics 15*(5), 533–551.

Karlsson, M., F. Klohn, and B. Rickayzen (2012). Are the dimensions of private information more multiple than expected? Information asymmetries in the market of supplementary private health insurance in England. *Cass Business School Working Paper 197*.

Karlsson, M., T. Nilsson, and S. Pichler (2014). The impact of the 1918 Spanish flu epidemic on economic performance in Sweden: an investigation into the consequences of an extraordinary mortality shock. *Journal of Health Economics 36*, 1–19.

Kiil, A. (2012). What characterises the privately insured in universal health care systems? A review of the empirical evidence. *Health Policy 106*(1), 60–75.

King, D. and E. Mossialos (2005). The determinants of private medical insurance prevalence in England, 1997–2000. *Health Services Research 40*(1), 195–212.

Klohn, F. and C. Strupat (2013). Crowding out of solidarity? Public health insurance versus informal transfer networks in Ghana. *Ruhr Economic Paper* (432).

Kobayashi, L. C., J. Wardle, and C. von Wagner (2014). Limited health literacy is a barrier to colorectal cancer screening in England: evidence from the English Longitudinal Study of Ageing. *Preventive Medicine 61*, 100–105.

Landmann, A., B. Vollan, and M. Frölich (2012). Insurance versus savings for the poor: why one should offer either both or none. *IZA Discussion Paper No. 6298*.

Lange, F. (2011). The role of education in complex health decisions: evidence from cancer screening. *Journal of Health Economics 30*(1), 43–54.

Larsson, K., I. Kåreholt, and M. Thorslund (2008). Care utilisation in the last years of life in relation to age and time to death: results from a Swedish urban population of the oldest old. *European Journal of Ageing 5*(4), 349–357.

Lechner, M. (2009). Long-run labour market and health effects of individual sports activities. *Journal of Health Economics 28*(4), 839–854.

Leider, S., M. M. Möbius, T. Rosenblat, and Q.-A. Do (2009). Directed altruism and enforced reciprocity in social networks. *The Quarterly Journal of Economics 124*(4), 1815–1851.

Leidl, R. (2008). A model to decompose the performance of supplementary private health insurance markets. *International Journal of Health Care Finance and Economics 8*(3), 193–208.

Leuven, E. and B. Sianesi (2014). Psmatch2: Stata module to perform full mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. *Statistical Software Components*.

Ligon, E. and L. Schechter (2012). Motives for sharing in social networks. *Journal of Development Economics 99*(1), 13–26.

Lin, W., Y. Liu, and J. Meng (2014). The crowding-out effect of formal insurance on informal risk sharing: an experimental study. *Games and Economic Behavior 86*, 184–211.

Maciosek, M. V., A. B. Coffield, N. M. Edwards, T. J. Flottemesch, M. J. Goodman, and L. I. Solberg (2006). Priorities among effective clinical preventive services: results of a systematic review and analysis. *American Journal of Preventive Medicine 31*(1), 52–61.

Marmot, M., D. Altman, D. Cameron, J. Dewar, S. Thompson, M. Wilcox, et al. (2013). The benefits and harms of breast cancer screening: an independent review. *British Journal of Cancer 108*(11), 2205–2240.

Martinez-Giralt, X. and P. Barros (2013). *Health economics: an industrial organization perspective*. Routledge.

Meier, V. and M. Werding (2010). Ageing and the welfare state: securing sustainability. *Oxford Review of Economic Policy 26*(4), 655–673.

Mensah, J., J. R. Oppong, and C. M. Schmidt (2010). Ghana's national health insurance scheme in the context of the health MDGs: an empirical evaluation using propensity score matching. *Health Economics 19*(1), 95–106.

Mocan, N. and D. T. Altindag (2014). Education, cognition, health knowledge, and health behavior. *The European Journal of Health Economics 15*(3), 265–279.

Morduch, J. (1999). Between the state and the market: Can informal insurance patch the safety net? *The World Bank Research Observer 14*(2), 187–207.

Morgan, S. L. and J. J. Todd (2008). A diagnostic routine for the detection of consequential heterogeneity of causal effects. *Sociological Methodology 38*(1), 231–281.

Mullainathan, S., J. Schwartzstein, and W. J. Congdon (2012). A reduced-form approach to behavioral public finance. *Annual Review of Economics 4*, 17–1.

Norton, E. (2000). Long-term care. *Handbook of Health Economics 1*, 955–994.

Nuscheler, R. and T. Knaus (2005). Risk selection in the German public health insurance system. *Health Economics 14*(12), 1253–1272.

Oldach, B. R. and M. L. Katz (2014). Health literacy and cancer screening: a systematic review. *Patient Education and Counseling 94*(2), 149–157.

Olivella, P. and M. Vera-Hernández (2013). Testing for asymmetric information in private health insurance. *The Economic Journal 123*(567), 96–130.

Oruč, N. et al. (2011). Do social transfers 'crowd-out' remittances: evidence from Bosnia. Technical report, The Vienna Institute for International Economic Studies.

Palangkaraya, A. and J. Yong (2009). Population ageing and its implications on aggregate health care demand: empirical evidence from 22 OECD countries. *International Journal of Health Care Finance and Economics 9*(4), 391–402.

Paraponaris, A., B. Davin, and P. Verger (2012). Formal and informal care for disabled elderly living in the community: an appraisal of French care composition and costs. *The European Journal of Health Economics 13*(3), 327–336.

Pavan, M. and A. Colussi (2008). Assessing the impact of public transfers on private risk sharing arrangements: evidence from a randomized experiment in Mexico. *University College Dublin. Geary Institute*.

Payne, G., A. Laporte, R. Deber, and P. Coyte (2007). Counting backward to health care's future: using time-to-death modeling to identify changes in end-of-life morbidity and the impact of aging on health care expenditures. *Milbank Quarterly 85*(2), 213–257.

Platteau, J.-P. (2000). *Institutions, social norms, and economic development*, Volume 1. Psychology Press.

Powell-Jackson, T., K. Hanson, C. J. Whitty, and E. K. Ansah (2014). Who benefits from free healthcare? Evidence from a randomized experiment in Ghana. *Journal of Development Economics 107*, 305–319.

Propper, C., H. Rees, and K. Green (2001). The demand for private medical insurance in the UK: a cohort analysis. *The Economic Journal 111*(471), 180–200.

Ravishankar, N., P. Gubbins, R. J. Cooley, K. Leach-Kemon, C. M. Michaud, D. T. Jamison, and C. J. Murray (2009). Financing of global health: tracking development assistance for health from 1990 to 2007. *The Lancet 373*(9681), 2113–2124.

Robins, J. M., A. Rotnitzky, and L. P. Zhao (1995). Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *Journal of the American Statistical Association 90*(429), 106–121.

Rothschild, M. and J. Stiglitz (1976). Equilibrium in competitive insurance markets: an essay on the economics of imperfect information. *The Quarterly Journal of Economics 90*(4), 629–649.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology 66*(5), 688.

Rubin, D. B. (1977). Assignment to treatment group on the basis of a covariate. *Journal of Educational and Behavioral statistics 2*(1), 1–26.

Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiology of Aging 30*(4), 507–514.

Schmitz, H. and M. Westphal (2015). Short- and medium-term effects of informal care provision on female caregivers' health. *Journal of Health Economics 42*, 174–185.

Seddoh, A., S. Adjei, and A. Nazzar (2011). Ghana's national health insurance scheme: views on progress, observations, and commentary. *Accra, Ghana: Centre for Health and Social Services 20*.

Serper, M., R. E. Patzer, L. M. Curtis, S. G. Smith, R. O'Conor, D. W. Baker, and M. S. Wolf (2014). Health literacy, cognitive ability, and functional health status among older adults. *Health Services Research 49*(4), 1249–1267.

Service, G. S. (2008). Ghana living standards survey, report of the fifth round (GLSS 5). Technical report.

Snelgrove, J. W., H. Pikhart, and M. Stafford (2009). A multilevel analysis of social capital and self-rated health: evidence from the British Household Panel Survey. *Social Science & Medicine 68*(11), 1993–2001.

Socialstyrelsen (2008). *Äldre – Vård och omsorg 2007*. Socialstyrelsen.

Steel, N., F. A. Huppert, B. McWilliams, and D. Melzer (2004). Physical and cognitive function. *Institute for Fiscal Studies, 2004*.

Townsend, R. M. (1994). Risk and insurance in village India. *Econometrica: Journal of the Econometric Society 62*(3), 539–591.

Tsai, M.-C. and D.-B. S. Dzorgbo (2012). Familial reciprocity and subjective well-being in Ghana. *Journal of Marriage and Family 74*(1), 215–228.

Udry, C. R. and T. G. Conley (2004). Social networks in Ghana. *Yale University Economic Growth Center Discussion Paper* (888).

United Nations (2013). *World Population Prospects: The 2012 Revision*. United Nations.

van Baal, P. H. and A. Wong (2012). Time to death and the forecasting of macro-level health care expenditures: some further considerations. *Journal of Health Economics 31*(6), 876–887.

van de Ven, W. P., K. Beck, C. Van de Voorde, J. Wasem, and I. Zmora (2007). Risk adjustment and risk selection in Europe: 6 years later. *Health Policy 83*(2), 162–179.

Wagstaff, A. (2010). Social health insurance reexamined. *Health Economics 19*(5), 503–517.

Wahhaj, Z. (2010). Social norms and individual savings in the context of informal insurance. *Journal of Economic Behavior & Organization 76*(3), 511–530.

Wallis, G. (2004). The determinants of demand for private medical insurance: evidence from the British Household Panel Survey. `http://repec.org/res2004/Wallis.pdf`.

Werblow, A., S. Felder, and P. Zweifel (2007). Population ageing and health care expenditure: a school of red herrings? *Health Economics 16*(10), 1109–1126.

Wilmoth, J., K. Andreev, D. Jdanov, and D. Glei (2007). Methods protocol for the human mortality database. *University of California, Berkeley, and Max Planck Institute for Demographic Research, Rostock. URL: http://mortality. org [version 31/05/2007]*.

Wilson, E. A., M. S. Wolf, L. M. Curtis, M. L. Clayman, K. A. Cameron, K. V. Eigen, and G. Makoul (2010). Literacy, cognitive ability, and the retention of health-related information about colorectal cancer screening. *Journal of Health Communication 15*(2), 116–125.

Wong, A., P. van Baal, H. Boshuizen, and J. Polder (2011). Exploring the influence of proximity to death on disease-specific hospital expenditures: a carpaccio of red herrings. *Health Economics 20*(4), 379–400.

Wooldridge, J. (2003). Introductory econometrics. *A Modern Approach. 2nd edition. Thomson South-Western*.

Wübker, A. (2012). Who gets a mammogram amongst European women aged 50-69 years? *Health Economics Review 2*(1), 1–13.

Wübker, A. (2014). Explaining variations in breast cancer screening across European countries. *The European Journal of Health Economics 15*(5), 497–514.

Yudkin, P. and I. Stratton (1996). How to deal with regression to the mean in intervention studies. *The Lancet 347*(8996), 241–243.

Zweifel, P., S. Felder, and M. Meiers (1999). Ageing of population and health care expenditure: a red herring? *Health Economics 8*(6), 485–496.