Essays on the Microfoundations of

Competition and Choice in Public Service Delivery

SEBASTIAN JILKE

Layout and printing: Optima Grafische Communicatie, Rotterdam, The Netherlands

Essays on the Microfoundations of Competition and Choice in Public Service Delivery

Essays over de microgrondslagen van de concurrentie en de keuze op publieke dienstverlening

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus

Prof.dr. H.A.P. Pols

and in accordance with the decision of the Doctorate Board.
The public defence shall be held on

Friday, the 20th of March 2014 at 11.30 hours

by

Sebastian Rainer Jilke
Born at Bad Soden am Taunus (Germany)

Doctoral committee

Promotors:	Prof.dr. S. Van de Walle
	Prof.dr. M. Haverland
Other members:	Prof.dr. W.J.M. Kickert
	Prof.dr. A.J. Steijn
	Prof.dr. G.G. Van Ryzin

ACKNOWLEDGEMENTS

It's finished, finally. But it takes some time to realise that all the blood, sweat and tears – and a good share of luck – finally got me to submit this PhD thesis. I naively assumed that the moment I submit it will be a life-changing experience – it was not. Yet, it should be celebrated nevertheless. Some weeks after my submission I felt relieved, but soon realised the work ahead of me. A PhD thesis is a good moment to look back in order to see what lies in front of you; we may have created some usable knowledge, but this tiny plant needs care.

I was not alone in working on this thesis. Many great friends and colleagues have contributed their fair share to this book, be it in the form of co-authoring one of the empirical chapters, providing comments, inspiring me, but also distracting me from work, or simply being there in times of ups and downs. You deserve more than a thank you; naming you hereafter is the least I can do.

First of all, I would like to thank my supervisors, Professors Steven Van de Walle and Markus Haverland. Steven, thank you for providing me the opportunity to do a PhD with you. You have not only been a great academic mentor to me, but were also extremely supportive with each new endeavour that I wanted to tackle – and there were many of them, but instead of forcing me to go back to my PhD you gave me the freedom and support I needed. The lessons I learned from you in many areas gave me a solid foundation for this thesis, and other work. Your enthusiasm for good research has always been a great source of inspiration. On a personal side, you have always been interested and caring, and fun to be with. Thanks for everything. Markus, although you joined-in during the second half of this dissertation, your support and rigorous reading of my works have been of great value to this dissertation. I admire your honest and straight-forward way of communication, which has made working with you a pleasure.

I also want to extend my gratitude to the other honourable members of the internal PhD committee, Professors Walter Kickert, Bram Steijn, and Gregg G. Van Ryzin. Walter, thank you for agreeing to sit on my PhD committee. You have been a great colleague to work with, and I would like to thank you in particular that you helped me in gaining relevant teaching experiences by letting me teach a part of your course on international public management – I learned some valuable lessons from you in this capacity. Bram, thank you for taking over the job as secretary of my PhD committee. While we work in slightly different areas of study, you always have been interested in my work, and provided very helpful comments on one of my papers

– it is now almost accepted, and your advice helped me a great deal to reshape it. Gregg, thank you for agreeing to sit on my internal committee. Through various projects we have worked more closely together in the past two years. During that time I learned a lot from you, be it in the area of experimentation, or more general in the field of citizen attitudes and behaviours. Your knowledge and experience in many areas is impressive, and it has always been fun to work with you.

Further, I would like to thank the honourable members of the external committee, Professors Gerhard Hammerschmid and Oliver James. Gerhard, through our joint involvement in the COCOPS project we have met each other regularly and your advices and comments on the research I am doing have always been of great value. You are always very kind and interested, I am happy that you accepted being part of my PhD committee. Oliver, in one way or another, you have been involved in my PhD work almost from day one. The advices and help I received from you were invaluable, and I learned lots from you. You are a great collaborator, and I admire your sincere interest in seeing credible research and answering relevant questions. And on a personal basis, you have been always great company. Thanks for being part of my PhD committee.

Next I would like to thank my paranymphes Dion Curry and Ziya Aliyev – and please be assured by this I do not mean nymph in a mythological way. Dion, besides being a great colleague you have become a friend to me. Our joint times at Erasmus and beyond have been great fun and you were always there to discuss important and unimportant matters. I am happy that you found a nice job at Swansea University, and look forward collaborating, or just having a couple of beers with you in the future. I am honoured that you agreed to become my paranymph. Ziya, we almost started our PhDs at the same time, and shared an office for over two years. It has been great fun having you as a roommate, and I particularly enjoyed our regular discussions about anything. But I also recall classical moments, for example when you were sunbathing in our office. I am honoured to have you as my paranymph.

I also would like to thank my collaborators on various projects for giving me the opportunity to learn from them, and for the help and advice I received. They include George Boyne, Stephan Grimmelikhuijsen, Asmus Leth Olsen, Carolyn Petersen and Lars Tummers. Special words of thanks go to my collaborator Rhys Andrews. Rhys, thanks a great deal for inviting me to Cardiff (and generously funding my stay) and being such a kind host and fun person to be with. In the first year of my PhD we started collaborating, and I learned a lot from you since. I admire how you manage doing cutting-edge research and participating in various academic debates at the same time, but also being such a genuine and kind person,

as well as great fun to have beers with. Oh, and I need to disappoint you, I still do not understand cricket rules. Another special thank you goes to my collaborator Marcos Fernández-Gutiérrez. I am truly grateful for your many helpful comments on various projects related to this thesis. You kicked-off the research line on citizen vulnerability in public infrastructure markets with your innovative PhD thesis, and I hope you find my results worthwhile. At last, I also want to thank Bart Meuleman for inviting me to KU Leuven and teaching me the hidden secrets of measurement. It was a truly intense week in Leuven, and I am happy that the results of our collaboration can be found in this dissertation.

Some words of thank you go also to my colleagues at Erasmus University Rotterdam, at the late Department of Public Administration. It has been a rough time since the reorganisation has hit the department. But I want to thank those who made it a nice and inspiring place to work – yet, some have already left the department. I especially want to thank Mansee Bal, Pablo Del Rio Loira, Stephan Dorsman, Harry Geerlings, Lasse Gerrits, Diana Giebels, Sebastiaan Goud, Sandra Groeneveld, Sanne Grotenbreg, Wouter Spekkink, Nadine van Engen, Ingmar van Meerkerk, Joris van der Voet, William Voorberg and Dirk Wolfson. A special thank you goes also to my room mates at T17-46. Moreover, I would like to thank the secretariat for great support in administrative matters: thank you Karin, and thank you Yneke. A final note of gratitude goes to the late Lalita – you were great. Also, I would like to thank Stefanie Gadella for her excellent job in formatting this thesis.

Outside Erasmus University, I would like to express my gratitude to a couple of colleagues who either commented on some parts of my dissertation, or with whom I had a good time at conferences. A warm thank you goes out to José-Manuel Alonso, Judith Clifton, Hans de Groot, György Hajnal (who I thank in particular for one of the funniest EGPA conference dinners ever), Christopher Pollitt, Soonhee Kim, Anca Oprisor, Riin Saavi, Vid Stimac, Wouter Van Dooren, and Sandra van Thiel. A special thank you goes to one of my early academic teachers Alexander Wegener – you were the one who convinced me to apply for the job at Erasmus, and you were right.

A special thing about the Netherlands is that many Dutchies are part of so-called clubs, most of the time book clubs. Here they meet to discuss books, obviously. I am proud that I was also part of such a club, yet, we mostly discussed our own research papers. A very warm thank you for excellent meetings, and great times with each other at various occasions goes to the PANDAs: Jan Boon, Sorin Dan, Chiara De Caluwé, Conny Hoffmann, Astrid Molenveld, Sjors Overman and Peter Oomsels.

My last words of thank you go to those close to me, my friends, family and my wife. You may have wondered on several occasions seeing me working late, and not coming out for drinks. Nevertheless, many of you have provided me with the needed distraction from work, and helped preventing me from becoming a man who has nothing else in life than his job. In particular I would like to thank my parents, my sister and my extended family for being there in times, but also for their understanding when I needed time for myself, and this thesis. Finally, I need to thank the most important person in my life, Eden. I am truly blessed having you by my side, and I enjoy every single minute of it. The important role you have played in finalising this thesis can not be expressed in words. I admire your kind personality, your grit and simply enjoy being around you. Thanks for being you, and thereby giving meaning to my life.

Sebastian Jilke
Kleve
January 2015

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction: The Microfoundations of Competition and Choice

1. Introduction: Choice and Competition in the Delivery of Public Services

The way how public services are delivered has changed fundamentally in past decades. While in the 1970s most public services such as energy, water or telecommunications were provided by state-owned monopolists, nowadays these services are delivered by a large array of different providers. In many countries public service providers operate alongside private companies. In Germany, for example, electricity services are frequently offered by municipally-owned public utilities who directly compete with private, often multinational, companies. This is clearly a situation in which public sector organisations have to compete with other suppliers in order to stay in business, and thus can no longer rest on their monopolist-status. The same holds true for many other different types of public services. Looking at the European public infrastructure sectors such as gas, post, rail, airlines, electricity, local transport and telecommunications, we can see that these markets experienced massive market deregulations in past decades in terms of entry regulations, market structure and public ownership status. This can be illustrated by looking at the OECD's market regulation indicators (Conway and Nicoletti, 2006). The main index[1] ranges from 0-6, and higher values indicate a high degree of market regulation (for example the presence of a national monopoly), while lower values stand for larger degrees of deregulation, including lower entry barriers for competitors, a lower degree of market concentration, and ownership diversity. Figure 1 displays these developments for a period of 32 years, including all mentioned public infrastructure sectors, through a single country measure. Indeed, with no exception, European countries that have been included in OECD's market regulation index experienced large deregulations of their public infrastructure markets.

What happened here? Two related trends can be observed that contributed to the deregulation of these markets. Firstly, the integration of the European Union into a common market and secondly, the changing ideology of policy-makers on the appropriate role of the state in society. The first trend, the European integration process and the creation of the single market, fostered the liberalisation – and privatisation – of public infrastructure services and made them subject to greater competition (Bognetti and Obermann, 2008; Clifton, Comin and Díaz-Fuentes, 2006; Prosser, 2005) – although the speed and depth of liberalisation varies across countries and service sectors (Conway, Janod and Nicoletti, 2005). Public infrastructure services were subject to liberalisation and deregulation reforms with

[1] The index is composed out of the following regulatory areas: barriers to entry, public ownership, market structure, and vertical integration (Conway and Nicoletti, 2006).

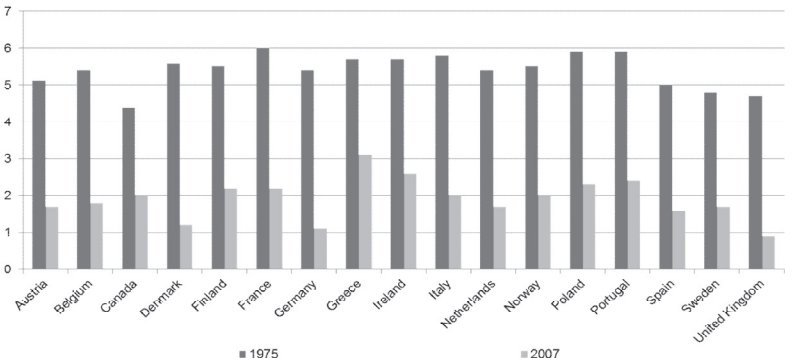

Figure 1: Market regulation in public infrastructure service sectors across the European OECD-member countries (1975-2007).
Source: Conway, Janod and Nicoletti (2005)

the aim of establishing a common market for services, where European providers could enter new markets with relatively little regulatory burden (Pelkmans, 2001; Prosser, 2005). The ultimate aim was to better integrate national economies into a single European market. At the time of the Single Market Programme (1987) and the Maastricht Treaty (1992), the major emphasis of European public officials was on how these service arrangements and the markets in which they operate could be remodelled to fit into the European single market. However, in the late 1990s these services were recognized of playing a major part in developing a European citizenship through their general-interest character (Clifton et al., 2013; Héritier, 2001; 2002; Prosser, 2005; Van de Walle, 2006; 2008).

The second related trend that contributed to the liberalisation of European public infrastructure services has a slightly different starting point, the advent of public choice theory in the 1960s (e.g. Arrow, 1963; Buchanan and Tullock, 1962). Inspired by neo-classical economic theory, public choice theorists emphasize (1) accountability deficiencies in the relationship between citizens, elected politicians and public servants (e.g. Moe, 1984; Tullock, 1965) and (2) the role of bureaucrats' self-interest (e.g. Downs, 1967; Niskanen, 1971), thereby justifying alternative ways of how bureaucracies should look like. For example, the budget-maximisation model of Niskanen (1971) argues that the fact that bureaucrats ultimately aim to maximise their own self-interests, would result in increasing their own office's budgets and authority[2]. Based on his analysis of the budget-maximising bureaucrat, it has been argued that governments should be smaller and controlled tightly, otherwise this

[2] Dunleavy (1991) later on developed the model of bureau-shaping which basically states that instead of maximizing their budgets, bureaucrats aim to shape their agencies in a way that it maximizes their

would lead to a too large and therefore inefficient public sector (Lane, 2000; see also Acemoglu and Verdier, 2000; Egeberg, 1995). Related to this idea was the conception of citizens' individual autonomy (Bognetti and Obermann, 2008). Citizens should not anymore passively receive the services that were provided by the state monopolist, but actively choose the offers (from competing providers) that best match their needs and demands. This shift from collective oriented models to individual customer based models emphasise the neoliberal imperative at the heart of market oriented reforms (Clarke and Newman, 2007; Learmonth and Harding, 2006). Politicians like Margaret Thatcher and Ronald Reagan were strongly influenced by neoclassical economic theory, and public choice theory in particular, and the related ideology of Neoliberalism, that is "*[…] the idea that the market offers the best solutions to social problems and that governments' attempted solutions, in contrast, are inefficient and antithetical to the value of freedom*" (Holland et al., 2007, p. xi). Or in other words, states are inefficient and markets can do better. The theoretical dogma behind this assumption is that competitive markets reach an equilibrium between citizens' needs and demands, and the actual quality and prices of the services that are offered. These markets seek to overcome the market failure situations that typically occur when public services are provided through a monopolistic provider, by establishing a market environment where, ideally, multiple service providers compete for customers (Savas, 1987; Ostrom and Ostrom, 1971).

On this basis, several countries enacted reforms of government modernisation, including introducing market-type mechanisms to the public sector (Hood, 1991; Pollitt and Bouckaert, 2011; Osborne and Gabler, 1992). Expressions of this trend were, among others, the introduction of competitive tendering regimes for the contracting-out of public service delivery (introduced the first time in 1978 in Arizona, Phoenix, see Osborne and Gabler, 1992), and later the marketisation, or liberalisation of public service markets (Bartlett and Le Grand, 1993; Florio, 2013) and thereby the introduction of choice into the delivery of public services (Le Grand, 2007).

Both reform trends gave rise to the introduction of market-elements into the provision of European public infrastructure services. Two elements are important considering here: competition and choice. Within public service markets, it is assumed that increasing competition between service providers would provide market incentives to produce better services for lower prices in order to sustain on the market. In other words, service providers would need to compete for customers. This logic implies that citizens in practice would have the choice to

personal utility (see James [2003] for an analysis of the UK's Next Steps reform initiative through the lenses of bureau-shaping theory).

exit service providers (Hirschman, 1970, pp. 21-25). Indeed, a key attribute in the provision of public infrastructure services is that the classical exit option of completely withdrawing from the service in question is often not feasible, too difficult or associated with extremely high costs (see Clifton et al., 2012). For instance, service users who choose to exit electricity or gas services would likely face severe effects on their individual wellbeing – especially in the winter. Here, choice, that is switching between different (public or private) service providers as a sub-form of exit (Dowding and John, 2012), becomes important[3]. Thus through switching, or implicitly threatening to switch via complaints (what Hirschman [1970] calls "voice"[4]), citizens send market signals to providers which are expected to create incentives to deliver greater value for money in order to keep existing customers, as well as attract new ones. As a results, it is assumed that a long-run equilibrium would be achieved between citizens' demands and preferences, and the price and quality of the offered services.

In this chapter we will review the microfoundations of competition and choice that underlie contemporary public service delivery arrangements and translate them into the research questions that guide this dissertation. Therefore, the next section will give an introduction into the microfoundations of competition and choice and distil two basic (and often implicit) assumptions that underlie the introduction of competition and choice into public service delivery. On this basis we will introduce an alternative theoretical view on citizens' market behaviours by drawing upon bounded rationality and decision-theory (section 3). In section 4 we will introduce our guiding research question that stems from the discussions in the preceding sections and operationalise it into four interrelated research questions. Section 5 introduces our empirical testing ground, public infrastructure services, and we will provide a reasoning why we chose this specific service sector. The following section will then discuss the theoretical and practical contributions of this study, as well as its value for the academic discipline of public administration as such (section 6). The final section of this chapter will provide an outline of the dissertation (section 7).

[3] Provider choice can be regarded as a result of competition, as a necessary but not sufficient condition for it. For competition to be translated into choice regimes, policies need to be enacted that allow citizens to effectively switch among providers. Thus choice is enhanced by regulatory policies and their interactions with citizens' market behaviour (Armstrong and Sappington, 2006; European Commission, 2012).

[4] Hirschman's definition of voice is broader than only encompassing complains, including also "*[...] general protest addressed to anyone who cares to listen*" (1970, p. 4). Dowding and John (2012) highlight that complaining to providers is indeed one of three sub-forms of how citizens can use their voice-option.

2. The Microfoundations of Competition and Choice

The above described neo-classical perspective on public service delivery under competition tends to assume that inducing competition into public services and increasing the number of service providers to choose from would result in an optimal allocation of available resources[5]. Or in other words, public services would become cheaper and better. This rests on an important set of micro-level assumptions, the microfoundations of competition and choice.

Microfoundations, defined as "*[…] the individual-level behaviour that underlies social activity*" (Stoker and Moseley, 2010, p. 8), are crucial for understanding how individuals actually tick and how these mechanisms influence aggregate level social outcomes. As we have seen in the previous part of this chapter, the theoretical accounts on the introduction of competition and choice into public service delivery make important assumptions about individual-level motivations and behaviours, in particular stemming from the idea that individuals are rational, utility-maximizing calculators (see also Stoker and Moseley, 2010). Indeed, macro-level reforms of introducing market elements into the delivery of public services are built around the conception of the *Homo Oeconomicus*. However, there exists limited awareness, and thus even less concrete empirical evidence whether these microfoundations of competition and choice get it right[6]. For this study we identify two important micro-level assumptions behind the introduction of competition and choice into public service delivery: The first on is that **citizens act as rational customers within liberalised public service markets** (*Homo Oeconomicus assumption*). The extension of citizen choice assumes that citizens identify and subsequently choose the service offer that best matches their needs and demands (European Commission, 2004; Stone, 2005). Or in other words, they reflexively choose the one service provider that provides them with the highest increase in their personal utility. Second, this also implies that **all citizens in their role as customers act equally rational.** This is what we will call the *homogenous rationality assumption*. It would mean that all citizens in their role as customers will take those decisions that are

[5] It has been argued that choice can indeed lead to a greater responsiveness to service users' needs and demands, but only if it meets two conditions: 1) competition must be real (there must be real competitors), and 2) there must be a real choice among alternatives (Le Grand, 2007).

[6] While there exist a large array of empirical studies in the behavioural sciences that demonstrate deviations from the conception of the rational man (see most prominently Kahneman and Tversky, 1979), to our knowledge, there exist limited works that test these assumptions in the context of liberalised public infrastructure services (for a set of studies on the potential consequences of non-rational market behaviour on vulnerability and citizen satisfaction with public infrastructure services, see Fernández-Gutiérrez, 2011).

close to their individual optimum, and hence increase their welfare homogenously[7]. Put differently, it is assumed that all citizens in their role as customers are equally able to take up the opportunities of increased competition and choice. Their behaviour as customers is equally rational for all groups of citizens. Thus, liberalisation reforms have been implemented by not only assuming that all groups of citizens act equally rational, but also that increases in individual welfare would be allocated evenly among them (European Commission, 2004; see also Clifton et al., 2011a).

3. Bounded Rationality and Insights from Decision Theory

As a response to theories of rational decision-making, which as we have seen form the microfoundations of competition and choice, Simon (1955) developed the concept of bounded rationality (see also Simon, 1972). This concept assumes that individual decision-making is not fully rational but is influenced by individual-level limitations, including uncertainty, cognitive constraints in processing information and information overload. Because of such personal limitations, people may end-up making poorer (non-optimal) decisions, including sticking with their default. In other words, according to the concept of bounded rationality, individuals' risk of making a poor decision increases when, for example, the number of options to choose from goes up – as is the case when the number of service providers increases. These assumptions correlate with findings from the decision-making literature on information overload, which suggest that as the amount of information to be processed grows, decision-making becomes poorer and also less likely (Chen et al., 2009; Hwang and Lin, 1999; Lee and Lee, 2004). This is mainly because individuals have limited capacities to deal with information when it comes to making decisions, and, when those limits are reached, individuals tend to become confused (Miller, 1956; Timmermanns, 1993). Consequently, if the number of options available goes beyond a certain limit, the risk of making a poor decision increases. As a result, the likelihood of staying with one's current service provider increases because this represents a safe haven, a so-called satisficing option – a situation which has been more generically described as 'status-quo bias' (Samuelson and Zeckerhauser, 1988). Indeed, studies in the field of applied psychology indicate that increasing the number of alternatives first results in a positive effect on customers' choice behaviour, but eventually the effect becomes negative (Botti and Iyengar,

[7] Hirschman (1970) – and later on the theory of marginal consumers (Schneider, Teske, and Marshall, 2002) – provides an alternative account to the claim that all layers of society will increase their welfare as a result of greater competition and choice. Here it is argued that for competition and choice to work it is sufficient to have certain amounts of alert consumers in markets that send market signals to providers. Those who are passive (the inert) will benefit from the trickle-down effects of the market signals of their more active counterparts. We will get back to this point in chapter 6.

2006, 2004; Iyengar and Lepper, 2000; Schwartz, 2005; Shah and Wolford, 2007), supporting the assumptions of information (and subsequently choice) overload.

Individual capabilities in processing information, and hence the propensity for taking a risk based on a possibly poor decision, has been found to vary largely among different socio-educational groups (Dohmen et al., 2007; Falch and Sangren, 2006; Hjorth and Fosgerau, 2010). That is, determining the best offer becomes more difficult for this particular group of people, and making a choice then represents a risk to them. This would mean that citizens who are potentially vulnerable as customers (Clifton et al., 2011a) (for example those with low levels of formal education versus those with higher levels[8]) would show a greater risk aversion and thus be less likely to switch among service providers, mainly because of their limited capacities in processing and evaluating necessary information. Against this background we question the identified microfoundations of competition and choice.

4. The Present Study: Research Questions and Overview

Within this study we examine the limitations of rational accounts about how citizens behave within public service markets. By drawing upon behavioural research on individual decision making we offer an alternative view of how citizens actually behave within liberalised public service markets and test it empirically. Concretely, we investigate whether citizens indeed make rational decisions in public service markets where competition and choice have been inserted. This leads to the central research question of this dissertation:

Do we find empirical support for the microfoundations of introducing competition and choice – that is assuming citizens acting as rational customers – within liberalised public infrastructure services?

Our guiding research question will be addressed in a comprehensive and integrated way by operationalising it into four interrelated research question (RQ). They are organised along the lines of the *Homo Oeconomicus assumption* (RQ 1), and the *homogenous rationality assumption* (RQ 2 and 3), as well one additional research

[8] In this study, we focus on people's educational attainment as a particular element of the concept of vulnerability, that is cognitive ability (more in chapters 4-6), as it largely affects citizens' resources for participating in the market (Hogg et al., 2007). One the one side education impacts the development of skills for customer empowerment, including the acquisition of information and the knowledge of how to interpret them. On the other side a low level of formal education is strongly associated with people's limited resources for processing and evaluating information, making in it an excellent candidate for examining the homogenous rationality assumption.

question that asks how to study these theoretical questions within a cross-national context (RQ 4).

1. Do citizens in their role as customers become more likely to switch away from their current public service provider after experiencing a service failure and when choice is increased?
2. Do different layers of society (particular those who are better versus less well educated) differ in their abilities to send markets signals to service providers in public service markets were choice has been implemented to varying degrees?
3. Are there differences in individual welfare (self-perceived affordability of services) between better and less well educated citizens in public service markets were competition and choice have been implemented to varying degrees?
4. How to study the microfoundations of competition and choice through survey data of citizen satisfaction with public infrastructure services, while simultaneously accounting for respondents' heterogeneous response behaviours across countries?

More precisely, in this study we investigate deviations from the mainstream model of citizens acting as rational customers. In doing so, we consecutively test a set of research questions that are of great theoretical and practical pedigree for the way of how contemporary public service delivery is organised. We first examine the Homo Oeconomicus assumption by asking whether increasing choice leads citizens to switch providers after experiencing poor services (*research question 1*). Next we address the homogenous rationality assumption by looking at whether different layers of society (particular those who are better versus less well educated) differ in their abilities to switch between service providers in public service markets were choice has been implemented to varying degrees, and whether we see similar patterns for their complaint behaviour (*research question 2*). If we indeed observe different ways of how citizens in their role as customers act within liberalised public service markets, then the degree to which they "benefit" from these market arrangements may also vary among different layers of society. Thus we will investigate whether disparities between better and less well educated service users are more, or less, strong in markets where choice and competition have been introduced to varying degrees *(research question 3*). At the end of the study we will also provide an methodological outlook of how to study the microfoundations of competition and choice through survey data of citizen satisfaction with public infrastructure services, while simultaneously accounting for respondents' heterogeneous response behaviours across countries (*research question 4*). Table 1 provides an overview how these research questions relate to the five empirical essays in this dissertation. All five essay have been presented at international conferences, and are either under

Table 1: Overview of Dissertation.

Chapter title	Research question	Publication status
1) Introduction	–	–
2) Citizens in charge? Reviewing the background and value of introducing choice and competition in public services	Literature review	Published in Edited Book
3) Public management reform and responses to decline in public services: An experimental evaluation of choice-overload	Homo Oeconomicus assumption (RQ1)	Revised and resubmitted to *Journal of Public Administration Research and Theory*
4) Choice and equality: Are vulnerable citizens worse-off after liberalisation reforms?	Homogenous rationality assumption (RQ2)	Published in *Public Administration*
5) Two-track public services? Citizens' voice behaviour towards liberalized services in the EU15	Homogenous rationality assumption (RQ2)	Published in *Public Management Review*
6) Vulnerable citizens in public service markets after regulatory reforms: Towards a price satisfaction gap?	Consequences of the homogenous rationality assumption (RQ3)	Revised and resubmitted to *Regulation & Governance*
7) We need to compare, but how? Measurement equivalence in comparative public administration	Methodological way forward (RQ4)	Published in *Public Administration Review*
8) Conclusions and discussion	–	–

review, accepted for publication, or already published in leading SSCI listed and international peer-reviewed journals. In addition, the literature review (chapter 2) has been published in an edited book.

Some of the empirical chapters have been developed with co-authors, thus we will subsequently describe the contributions that the author of this dissertation made to each. Chapter 2 (Tummers, Jilke and Van de Walle, 2013) was developed jointly, and the author of this dissertation took the lead in reviewing the literature and writing up the parts that concerned the logic behind the market mechanisms of public service delivery and the effects of competition and choice on citizens (overall contribution 45%). In chapter 3 (Jilke, Van Ryzin and Van de Walle, 2014), the author of this dissertation took the lead for all sections of the study (overall contribution 90%). Chapter 4 is single authored (Jilke, 2014). Chapter 5 (Jilke and Van de Walle, 2012) was led by the author of this dissertation (overall contribution 90%). In chapter 6 (Fernández-Gutiérrez, James and Jilke, 2013; alphabetical authorship), the author of this PhD thesis took the lead in the empirical analysis and writing-up of the results, as well as jointly worked on the theory section, the development of the hypotheses, and the conclusions (overall contribution 45%). In the final empirical

chapter (chapter 7) (Jilke, Meuleman and Van de Walle, 2014), the author of this dissertation took the lead for all sections of the study (overall contribution 90%).

5. Why Studying Public Infrastructure Services?

This study focuses on the introduction of competition and choice within the European[9] energy and telecommunication sectors, using these sectors as an empirical testing ground for our theoretical predictions. The obvious question that arises from this choice may be why not other sectors, such as health, or education? Indeed, competition and citizen choice have been introduced not only in the public infrastructure sectors, but are also visible in quasi-markets of public services such as healthcare, or education (Godwin and Kemerer, 2002; Le Grand, 2007; Schneider, Teske and Marshall, 2002). The National Health Service (NHS) in the United Kingdom (UK) is a prime example of the introduction of competition and choice into public service delivery (e.g. Bennett and Ferlie, 1996; Proper, Wilson and Burgess, 2006), but also the area of secondary education has a longstanding history of citizen choice mechanisms in many countries (e.g. Musset, 2012). However, we have to note that many of these services still operate under so-called quasi-market arrangements (Bartlett and Le Grand, 1993). In contrast, European public infrastructure services have been liberalised to much greater extent – yet, still displaying sufficient cross-national variation. Precisely for this reason, prominent scholars have called these sectors as an optimal laboratory for the study of liberalisation reforms (e.g. Clifton et al., 2011a; Florio, 2013; Milward, 2008). Moreover, European education and health systems differ substantially in terms of their institutional designs when it comes to the introduction of competition and choice (Musset, 2012; Wendt, 2009). Thus, comparing them across different countries may be of interest from an institutional point-of-view, but for examining the associated microfoundations and consequences of the introduction of choice and competition, European public infrastructure services provide a better-suited empirical testing ground for our theoretical predictions.

But can infrastructure services such as electricity and telecommunications really be regarded as public services, given that especially in liberalised markets they are increasingly delivered by non-public providers? In Europe there exist the legal term *services of general economic interest* which was introduced by the European Commission to replace varying national definitions of public services (Prosser, 2005). Services were labelled as being in the general interest if they are "*[...] subject to*

[9] With the exception of chapter 3, were the United States' (US) energy sector has been taken as an empirical case for the experimental design.

specific public service obligations by virtue of a general interest criterion" (Commission of the European Communities, 2004, citied after Van de Walle, 2008, p. 7). Thus the term encompasses services of an economic nature subject to competition and market rules, but also to public service obligations under a general-interest criterion due to its essential character for citizens (European Commission, 2003; Bauby, 2008; Clifton, Comín and Díaz-Fuentes, 2005). Thus they are not understood as being public in terms of their mere ownership status, but in terms of the objectives inherent to their supply, that is that the delivery of these services is in the public interest. With this approach, the EU officially recognises the key role of these services in strengthening solidarity, equity and social cohesion, as one of the cornerstones of the European social model (European Commission, 2004). Moreover, it also needs to be noted that what constitutes a public service or good (apart from Samuelson's [1954] narrow microeconomic definition of public goods) has in the past often been the result of policy choices[10] (see Kaul et al., 2003). Within this study, we concentrate on those services of general economic interest that are subject to general interest and therefore considered appropriate for delivery through market-based mechanisms but, at the same time, should also be made accessible to all layers of society – which constitutes their public service character. Here, electricity and telecommunication services are prime examples.

6. The Contributions of this Dissertation

This dissertation challenges two basic assumptions that underlie the introduction of competition and choice into the delivery of public infrastructure services. Until recently, few public administration scholars have emphasised potential flaws of the theoretical foundations that form the basis of market-oriented reforms in public service delivery (e.g. Clifton et al., 2011a; Stoker and Moseley 2010), yet even fewer have engaged in the empirical testing of these microfoundations. Thus the major theoretical contribution of this dissertation is twofold. First, this dissertation identifies the key theoretical mechanisms at the micro-level that underlie the introduction of competition and choice into public service delivery. It contrasts these microfoundations with theoretical accounts from decision theory that suggest individual bounded rationality and choice-overload, thereby develop-

[10] See most prominently the example of access to clean water which is nowadays (mostly) considered a public good. However, in the 1800s access to clean water has been mainly seen as a private good being in the individual responsibility of the people. This eventually changed after the negative externalities of no access to clean water became public (e.g. the spread of pest and cholera; see for example John Snow's [1855] classical experiment on the cholera epidemic in London); here local authorities later on explicitly provided access to clean water (and sanitation) to their populations, making it a public service (Jilke, 2009, p. 75).

ing innovative theoretical predictions of how citizens in their role as customers behave within liberalised – and thus increasing complex – public service markets. Second, especially the theory of choice-overload has been subject to empirical testing within consumer research and applied psychology; this study is the first that empirically applies and extends this theoretical account to public services. Thereby, it not only provides useful insights of how people deviate from the classical model of rational decision-making, but also provides a valuable and innovative extension of the theory by considering service failures and subsequent citizen dissatisfaction, instead of simple consumption decisions (whether or not to buy a jam) – this extension will come helpful also in the field of applied psychology, where only recently calls for further investigations of the boundary conditions of choice-overload have been made (Cherney, Böckenholt and Goodman, 2010; Scheibehenne, Greifeneder and Todd, 2010)

On a more practical side, this dissertation contributes to a recent trend in evidence-based policy-making (e.g. Cartwright and Hardie, 2012; Davies, Nutley and Smith, 2000) by suggesting to overcome ideological reform dogmas of how the state *should* look like. In contrast, we provide evidence that ideology and empirical evidence can dramatically differ, and thus emphasise the need for credible empirical testing of what actually works, and why. More concretely, using a robust empirical strategy we show that, indeed, citizens do not behave in the same rational manner as policy-makers would have predicted. This provides food-for-thought of how contemporary designs for the delivery and especially regulation of public services could look like. This is especially important when considering the general interest character of public infrastructure services.

In terms of the academic field of public administration this dissertation makes also makes two important contributions to the discipline as such. The first is methodological innovation, and the second is the outline of a behavioural public administration. First, this study provides an example of using various innovative research techniques, such as experiments, multilevel analyses, and cross-national measurement techniques (such as multilevel item response theory mixture modelling) to study an important problem of theoretical and practical pedigree. Especially chapter 7 provides guidance of how cross-national researchers in the discipline can engage in cross-national analyses in a most robust manner. These research methodologies have received little or no attention in the field of public administration. Only recently the field has opened up to more robust research designs, but there is still a long way ahead of us. This dissertation contributes to this development. Secondly, our study is truly interdisciplinary in its nature, spanning over the disciplines of public administration, psychology, and behavioural economics. Thus it fits into

what Nobel laureate Daniel Kahneman (2013) has termed an applied behavioural science – making the processes within public administrations and public policy the unit of analysis of the behavioural sciences. Indeed, the study of the behavioural and decision-making components of public administration is something that eminent scholars such as Nobel laureate Herbert Simon have propagated in the early days of the discipline (Simon, 1955; see also Olsen, 2015). In this dissertation we lay another stepping stone for the interdisciplinary analyses of public administration theory from the micro-perspective of individual behaviour and evaluative judgements by drawing upon recent advances in our understanding of the psychology and behaviour of individuals that is the study of a *behavioural public administration* that examines the behavioural foundations of public administration theories.

7. Outline

In this dissertation we report from a series of empirical essays that investigate the microfoundations of competition and choice in public infrastructure services. With the exception of chapter 3, all of these empirical chapters examine European infrastructure sectors from a comparative perspective. Thus the European common infrastructure market provides us with an empirical testing ground for our theoretical predictions.

In chapter 2, we review the literature on competition and choice in public service delivery and extract theoretical and practical opportunities, and pitfalls that may arise after public service markets have engaged in opening up for competition and choice. This chapter sets the scene for the following empirical essays, starting with chapter 3. The third chapter first develops a theoretical model of choice overload as a response to service failure, and tests it empirically by means of a population-based survey experiment in the US electricity sector. Our experiment reveals that increasing the number of service providers *reduces* people's motivation to switch away from poor performing services. These findings stands in stark contrast to basic assumptions of rational decision-making, and thus provide robust evidence against the Homo Oeconomicus assumption. Moreover, these results are replicated with a second independent experiment, further increasing the external validity of our findings.

Chapter 4 and 5 look at whether citizens who are potentially vulnerable in their role as customers (in particular the less well educated) are less likely to send market signals to providers when compared to their better-off counterparts, thereby examining the homogenous rationality assumption. Here we analyse comparative survey data from a national representative sample of citizens' self-reported switching and

complaint behaviours across up to 25 European countries. We find that the less well educated are indeed significantly less likely to switch between providers, and also that they less frequently complaint. Given that previous studies have reported that especially this group of citizens is less satisfied with their services, this is an important finding. We, more importantly, test whether this gap is wider in markets with a lot of choice. In strongly liberalised sectors (such as mobile telecommunications), we find that the gap in switching between service providers is significantly stronger when the number of available service providers is large. Thus our findings provide no clear support for one of the basic theoretical foundations behind the introduction of competition and choice into public services, that is the homogenous rationality assumption. Instead, the opposite seems to hold true.

Chapter 6 investigates the consequences of this non-homogenous market behaviour by asking whether the self-perceived affordability of European public infrastructure services vary across socio-educational groups, and whether this gap is of stronger magnitude in markets where competition and choice have been implemented to greater degrees. We find that less well educated citizens are less likely to perceive their services as affordable when compared to the better educated, even after controlling for a number of potential socio-economic confounders. Moreover, this relationship seems to be unaffected by the degree of competition within national markets. However, when there is a higher national switching rate in markets, the self-reported welfare-gap between socio-educational groups is significantly smaller and eventually diminishes. This has been found in both, the electricity and telecommunications sector. We draw upon the theory of marginal consumers to explain these interesting results, suggesting that while we find some evidence against the microfoundations of competition and choice (i.e. citizens do not necessarily behave as rational decision-makers), markets seem to work (i.e. citizen's needs and preferences more closely match the services they receive) when there exists a critical mass of citizens that switch between providers, thereby providing sufficient market pressures.

Finally, chapter 7 shows how to study the microfoundations of competition and choice through survey data of citizen satisfaction with public infrastructure services, by introducing two important and innovative methodological techniques for dealing with respondents' heterogeneous response behaviour. Hence this chapter mainly contributes to the methodological innovation of examining the microfoundations of macro theories via comparative survey data.

Chapter 8 discusses the practical and theoretical implications for our research questions in the light of the empirical findings from previous chapters.

CHAPTER 2

Citizens in Charge? Reviewing the Background and Value of Introducing Choice and Competition in Public Services

This chapter has also appeared as Tummers, Jilke and Van de Walle (2013).

1. Introduction

Marketisation, the process of integrating market elements into the public sector, has been one of the core objectives of public management reform in many countries (Pollitt and Bouckaert, 2004). Offering choice between providers to citizens is an essential element in this marketisation. It is assumed that by introducing choice there will be less need for hierarchical steering. Instead of this, citizens would act as customers and send market signals to providers through either exercising choice or expressing voice. As a result, a better match between client preferences and services offered would emerge (Van de Walle, 2010). Based on this logic, reforms aimed at increasing choice opportunities have been introduced in various countries.

Introducing choice in public services was supposed to put citizens in the "driver's seat", making them in charge of their service provision (Kremer, 2006, p. 385). Many scholars have argued, however, that introducing client choice in public services may have had unintended negative effects, and that citizens do not always act as empowered public service customers (Hsieh and Urquiola, 2003; Van de Walle and Roberts, 2008; Wilson and Price, 2010).

This chapter provides an overview of the background, facilitators and pitfalls of introducing client choice and competition. To date, there have been a number of studies on introducing choice in specific sectors, such as healthcare (Glasby and Littlechild, 2009), utilities (De Bruijn & Dicke, 2006) and education (Teske and Schneider, 2001). Next to this, general works about choice have been written, which are often normative being either somewhat pro-choice (Le Grand, 2007) or more critical (Dowding, 1992; Greener, Simmons, and Powell, 2009). This chapter aims to provide a balanced overview view on choice examining its facilitators and its pitfalls. Throughout the chapter we will illustrate the background, facilitators and pitfalls with real-life examples from choice-related innovations in various countries, and examine available evidence on their effectiveness. In this way, we aim to increase our understanding of choice in various countries and sectors.

We start by providing a background of choice (Section 2). In this section, we will first focus on the important question: what is choice and how does it work? Following this, we will discuss the reasons why the choice movement has become so influential in public services in various countries. Section 3 then discusses potential facilitators of choice, such as the need for a good functioning market of demand and supply. Despite the high-minded rhetoric about choice it does not always function as desired. In Section 4, we therefore address some potential pitfalls of choice, such as the games played by suppliers, and increased disadvantages for citizens from

lower socioeconomic classes. The chapter ends with a conclusion on the values of choice in public services.

2. Background of Choice

Exit, voice and loyalty

Before going into the facilitators and pitfalls of choice, we discuss what is meant by the term 'choice', and the related term 'voice'. This discussion draws on the classical Exit-Voice-Loyalty framework of Albert O. Hirschman (1970). Hirschman developed this framework in order to understand the decline of public organisations, private organisations and states. According to him, there are basically two options when you are unsatisfied with a situation: you can either leave the situation, choice/exit option, or you can attempt to repair the situation, voice option. The degree of loyalty can influence choosing for exit or voice. Imagine for instance your relationship with a particular postal service organisation in a liberalised postal market. When you experience the costs and speed of their postal delivery as unsatisfying you can either stop using the services of the organisation and go to a competitor, or you can express your concerns to this organisation, in order to improve the situation. When you are loyal to the postal service organisation the voice option might be more rewarding in first instance than the more definitive exit option.

The distinction between choice (or exit) and voice can be applied to public services (Le Grand, 2007; Greener, Simmons and Powell, 2009; SIX, 2003). Here, choice often means that dissatisfied service users will opt out and move to providers perceived as better performing. In this way, providers have a strong incentive to deliver more value for money in order to keep their customers and attract new ones. The second mechanism is 'voice'. Dissatisfied citizens will express their discontent, and in this way will force providers to improve on aspects of service delivery. When the public service provider is a monopolist, voice will often be the only viable option given moving out of the country or stopping to use the service are often impossible.

In public services, we can differentiate between three different types of choice (see Dowding and John, 2007). 1) Physical relocation, or Tiebout exit; 2) switching between private and public providers; and 3) switching among public providers only. Firstly, a Tiebout exit occurs when service users purposefully change their residence in order to receive better public services, or lower taxes (Tiebout, 1956). One popular example is parents who move to a new house in a different school district to give their children a better quality education. Another, very recent, example is Gérard Depardieu, one of France's best-known actors, who is leaving France in

order to evade the high income taxes. The second option is to switch between private and public providers. This happens, for instance, in healthcare, when patients leave a public hospital for a private hospital or clinic expecting for instance better healthcare, or the same healthcare quality without waiting lists. Another example might be parents who send their children to private schools or universities (such as Harvard or Yale), instead of the (cheaper, but less prestigious) public institutions. The third choice option is to choose between various public providers. Choosing solely between public providers may be the case when citizens switch between different public schools, or health services. Citizens then stay within the public system of service provision. In this chapter we will focus on all three types of client choice.

Introducing choice can facilitate voice

Next to choice, citizens can also use the voice option by expressing their discontent (Jilke and Van de Walle, 2012). When choice is introduced, voice can become more important. Prior to marketisation of public services such voices were often unheard or ignored. Providers did not have a strong incentive to react, as citizens could not move to another provider. Voice allows providers to anticipate future exits and change service levels accordingly. Voice, in this way, is an early warning of exit, or even a threat of exit. Failure to deal with voice means providers will be confronted with a loss of their clients. Since funds now tend to follow clients rather than being paid as a lump-sum to providers, service providers will try to take the needs of their clients into account in order to stay in business. Thus choice and voice send signals to providers that complement each other in improving public service performance. As service users are prone to choice and voice, the autonomy for choosing resides then with the service users themselves and no longer with the provider (Wilson and Price, 2010; Clarke et al., 2007).

Introducing choice in public services

The former two sections introduced the related concepts of choice and voice. A number of countries have enacted reforms aimed at increasing the choice opportunities of citizens (Dowding, 1992; SIX, 2003; Fotaki et al., 2008). In order to understand why this happened, we will provide a short background of the introduction of choice in public services.

The introduction of reforms focused on increasing choice can be linked to the development of the public choice field of economics that started in the 1940s-1950s (Hayek, 1944; see for instance Black, 1948; Friedman, 1955). Public choice scholars analysed the behaviour of civil servants and politicians in public decision-making (Buchanan and Tullock, 1962). For instance, Downs (1967) looked at the behaviour of civil servants in his book "Inside bureaucracy". He noted that "[…] every official

acts at least partly in his own self-interest, and some officials are motivated solely by their own self-interest" (1967: p. 83). Based on the motives of the civil servants, he developed various ideal types, ranging from purely self-interested civil servants, motivated almost entirely by goals that benefit themselves, to statesmen, loyal to society as a whole. Related to this, Niskanen (1971) developed the 'budget maximising model'. Using this model he argued that bureaucrats ultimately aim to maximise their own self-interests, which results in maximising their agency's budgets and authority.

Although the field of public choice analyses what governments do (descriptive analysis, 'what is'), public choice scholars are also often concerned about what governments should look like (normative, 'what ought to be'). For instance, based on the work of Downs and Niskanen, it has been argued that governments should be small and controlled tightly. This is because the behaviour of civil servants ultimately leads to a public sector, which is too large and therefore inefficient (Lane, 2000). Furthermore, public choice scholars noted that as many civil servants were self-interested, they were less occupied with the interests of the citizens they were supposed to serve (Egeberg, 1995; Acemoglu and Verdier, 2000). Given this situation, it was argued, power should shift from civil servants to citizens wherever possible and introducing provider choice was seen as a proper way make this transition/shift.

Hayek, who can be considered to belong to the Austrian school of public choice (McNutt, 2002), also developed views about how the government should look. He was very critical about the role of the government. In his seminal work "Road to Serfdom" (1944), he argued that all forms of collectivism, government control of economic decision making through central planning, ultimately leads to tyranny. He noted that central planning is an inferior method of regulation given it's ineffectiveness due to being carried out by a limited number of people who possess limited information. Furthermore, it is undemocratic given the will of a small minority of people in power is imposed upon a large group of relatively powerless citizens. Hayek, strongly opposing Keynes, claimed that governments should have only a very limited role in the society and should only intervene when markets fail, such as in the case of negative externalities, the classic case being the factory which pollutes the environment.

The political (mis)use of public choice theory started roughly in the 1970s and 1980s, when economic crises and the collapse of the Communist bloc fuelled political opposition to state interventionism in favour of free market reform (Tummers, Bekkers and Steijn, 2012). Politicians like Margaret Thatcher in the U.K. and Ronald

Reagan in the U.S. were heavily influenced by public choice theory and the related ideology of neoliberalism. Neoliberalism is "[…] the idea that the market offers the best solutions to social problems and that governments' attempted solutions, in contrast, are inefficient and antithetical to the value of freedom" (Holland et al. 2007, xi). Based on the ideology of neoliberalism, several countries enacted reforms for the modernisation of government, such as introducing choice by denationalisation, disaggregation of public-sector units, and more explicit performance measures (Clarke and Newman, 1997; Harvey, 2007). In these ways, the ideology of neoliberalism, combined with the introduction of business-type managerialism, led to a number of public sector reforms under the label 'New Public Management' (NPM) (Pollitt and Bouckaert, 2004; Hood, 1991; Savas, 2000; Osborne and Gaebler, 1992).

One of the core New Public Management reforms focuses on the introduction of choice in public services (Dowding, 1992; SIX, 2003; Fotaki et al,. 2008; Hood, 1991). The introduction of choice into public service delivery aims at remedying the undesirable effects of state monopoly provided public services (Le Grand, 2007; Ostrom and Ostrom, 1971). This is done via the introduction of (quasi)-markets into public service delivery where providers compete for customers (Bartlett and Le Grand, 1993). There are many examples of the introduction of choice in public services, but possibly the most visible change is seen in the utility sectors (electric, gas and water services) where monopolised provision has been replaced by a system in which many providers compete for customers (De Bruijn and Dicke, 2006). Other sectors where choice has been introduced include healthcare (Glendinning, 2009; Tummers, Steijn and Bekkers, 2012), education (Godwin and Kemerer, 2002), social security (Sol and Westerveld, 2007), or postal services (Schulten, Brandt and Hermann, 2008).

In sum, it is clear that the notion of choice has become an important aspect in public services. Hence, it is of paramount importance that we gain an understanding of this phenomenon. In the remainder of this chapter, we will first concentrate on the structural conditions, or facilitators, for choice to work as intended. We subsequently address some pitfalls of introducing choice in public services.

3. Facilitators of Choice

For choice to function, citizens need to be able to act as consumers. This requires the presence of a functioning market, market information in order to make informed choices, and the presence of payment tools. This may require governmental inter-

vention to create market information, and to determine how citizens pay for the services (funding mechanisms), especially in a situation where public services have recently moved from monopolistic public provision to a quasi or full market. We label these 'facilitators of choice': they make choice possible. We will focus on three important facilitators: 1) market making, making sure enough providers are present, 2) providing market and service information, and 3) providing funding mechanisms.

Market making

Moving from monopolistic government-led service provision to a market for public services in which citizens can exercise choice requires that several suppliers actually enter the market. Without more than one supplier, marketisation and (semi-)privatisation will by definition not lead to more choice, but instead to a new dependence on one monopolistic provider.

In some sectors, there is substantial competition available. For instance, in an OECD paper of 2006, it was shown that the energy, transport and communications industries have become more opened to market mechanisms by reducing, among others, price controls and entry barriers (Conway and Nicoletti, 2006). However, there were substantial differences among countries. English-speaking countries and Germany had relatively open markets, while markets for energy, transport, and communications were more adverse to competition in France, Ireland and Greece.

When governments try to open up markets for choice and competition, two main approaches can be distinguished: a) taking measures to stimulate the emergence of new providers and b) protecting the market against predatory practices and market concentration. First, we will examine measures to stimulate the emergence of new providers. When a public service market is opened for competition this means that most service providers have to start anew. Such new providers may be former non-marketized services, or collaborations of former employees of such services. Working in a new market is a risky undertaking. This also explains the rise of large multinational (public) service provision conglomerates; they have specific competences and skills to open for business in newly opened areas, as well as sufficient capital and resources to carry the burden and risk of these new enterprises (Clifton and Díaz-Fuentes, 2008). In order to offer choice, governments need to intervene to make sure that the new market will actually have a sufficient number of providers, rising from a monopoly, via an oligopoly to possibly a near-perfect market. This can be done in various ways (Savas, 2000). One is through good management of the transfer of public companies or assets to the market, through privatising them, either as a whole, or through splitting them into separate lots. Alternatively, this can be done via supporting providers that want to start from scratch, for instance by

establishing training schemes, attractive legal and fiscal conditions, or investment support.

A second market making task for government is to protect new markets against abuse and market concentration. This is generally done through governmental regulation. In the utility sector, universal service obligations are a typical measure taken by governments to make sure the market will function as intended. Just as is the case in many other markets, governments also typically intervene in public service markets to avoid concentration, through, for instance, establishing specific regulatory bodies, or by strengthening competition authorities. In these ways, governments can make use of market making to facilitate choice in public services.

Providing market and service information

The second intervention governments can apply in order to increase choice opportunities, is to assist citizens in exercising choice and especially, assist them in making the 'right' choices by providing them with easily accessible and clear information on different service offers. Access to objective measures of service providers' performance supports not only service users' choices, but also provides an incentive to providers to improve the quality of their services (Le Grand, 2007). The same holds true for comparisons of prices. Citizens are expected to act as customers and use performance information to guide their choice of service providers (Coe and Brunet, 2006). Examples include league tables of school performance, and published performance data of hospitals, including waiting times for certain types of surgeries, or even mortality rates.

Examples of providing market and service information abound. For instance, from the U.K., uSwitch compares gas and electricity suppliers. NHS choices facilitates comparing service offers in health care. Another example from the U.K. healthcare are Patient Care Advisors (PCA). They not only act as suppliers of relevant information, but also give case specific advice, as well as help clients in making the necessary organisational arrangements with hospitals and other service providers. Recent evaluations on the use of PCAs indicate that they are highly regarded and frequently used by patients (Coulter, Le Maistre and Henderson, 2005). Furthermore, in the Netherlands, the website www.kiesBeter.nl ('choose Better') provides information on healthcare providers. On this website, it is noted that it "[…] is designed for all adult residents of the Netherlands who have questions in the field of healthcare, health insurance and health. The information on kiesBeter.nl is reliable and can help make choices in this area." The National Institute for Public Health and the Environment, part of the Ministry of Health, Welfare and Sports, developed this website. Hence, in this way, the Dutch government aims to provide

better market and service information in Dutch healthcare. More in general, Damman and Rademakers (2008) analysed over 50 websites from different countries concerning choice information for customers in healthcare. They noted that many countries do indeed provide such websites and that countries like the U.K. and the Netherlands are frontrunners in this respect.

Direct funding mechanisms: Money follows clients

Introducing choice also requires a different way of funding service providers. Traditionally, public service providers have been funded through lump-sum funding. As a result of further marketisation, funding has become increasingly tied to client numbers. An even more fundamental change related to the introduction of more provider choice has been to transfer funds directly to citizens. This facilitates choice opportunities for clients as they now have more power to choose. According to Baxter et al. (2011: p. 91), the aim of transferring funds directly to clients is "[…] to move away from service-led arrangements and give users more direct control over the resources available to them, so that services can be better tailored to their individual needs and circumstances." Providing citizens with vouchers or budgets can lead to 'real' empowerment (Morris, 1993). Giving citizens budgets can be seen as one of the most extreme forms of choice. It can mean choice on multiple dimensions, such as choice of location (where), choice of professional (who), choice of service (what) and choice of time (when).

In various sectors, experiments have been set up to transfer funds directly to clients, most often via vouchers or via direct budgets. Considering social security, Sol and Westerveld (2007) note that reintegration services often provide job seekers a grant or, more indirectly, a voucher, which can be cashed at the counter of various service providers. This incentivises service providers to improve their services. In the U.K., people living in so-called 'Employment Zones', areas with high long term unemployment, are able to receive a direct budget to set up in business, improve their skills or even buy clothes for a job interview. These Zones are managed by the Department for Work and Pensions. The Netherlands also experimented with reintegration budgets for partly disabled people. Here, this particular group of unemployed people could develop their own reintegration plan and make decisions regarding the reintegration companies from whom they wanted to purchase activation services (Van Berkel and Van der Aa, 2005). Finally, Germany also introduced placement and training vouchers for job seekers (Sol and Westerveld, 2007).

Next to social security, vouchers systems are also often used for school choice. This has been initiated by the work of Milton Friedman (1955). He wrote a seminal essay on the role of the government in education, in which he argued in favour of the use

of school vouchers. Based on, among others, the pioneering work of Friedman, a large body of literature developed concerning the advantages and disadvantages of using school vouchers (Godwin and Kemerer, 2002). In the United States, there is a large market concerning school vouchers, or scholarships, for private schools, and evidence suggests small positive achievement gains for students (Howell et al. 2002; Mayer et al., 2002; Rouse, 1998). On the other hand, evidence from other countries, such as Chile and New-Zealand, suggests that school vouchers have only limited positive effects, and can even have substantial negative side effects, such as harming disadvantaged students and low-income families (Hsieh and Urquiola, 2003; Ladd, 2002).

A final example of the use of direct funding mechanisms is the introduction of personal care budgets in home care. Personal Care Budgets give citizens money directly to pay for their own homecare, rather than the traditional route of providing services through regional health insurance carriers. Users are, for example, citizens with physical and sensory impairments or parents of disabled children. The U.K. was one of the first countries to introduce so-called 'cash-for-care' schemes (Glasby and Littlechild, 2009). Following the British example, many countries introduced Personal Care Budgets, such as France, Germany, Finland and Australia (for an overview, see Lundsgaard, 2005). These budget schemes all compensate care financially, aiming to give a stronger 'voice' to the client. The following quotation by a care client illustrates a vivid example of this (cited in Ungerson, 2004: p. 203):

"I mean, we have to have these carers and it's better than having social services that come in at a certain time and treat you like you're robots – you get up at a certain time, go to bed at a certain time and you function at a certain time. Whereas [with] your own carers, to a certain extent you have got control of what time you want to get up, what time you go to bed, things like that."

However, giving citizens budgets, and choice more in general, can also have substantial pitfalls. In the next section, we will focus on these pitfalls of choice.

4. Pitfalls of Choice

Choice does not always function as desired (see also Savas, 2000). We discuss and present empirical evidence of: 1) too much market power; 2) increasing inequality among citizens; 3) problems with using performance information; and 4) worsening work conditions. These pitfalls are related to the facilitators of choice.

High market power

Dowding (1992) argued that at least two positive alternatives are required for choice to be meaningful: a client should be able to choose between minimally a and b. For instance, you should be able to choose between going to a school which is close by or a school, which is further away, but with a better reputation. This is completely different from a negative choice between a or not-a; such as going to a school which is close by or not going to a school at all. This condition is not always fulfilled. One important situation in which this condition is not fulfilled is when there is high market power by providers. Market power can therefore be seen as an important pitfall when introducing choice (Baxter, Glendinning and Greener, 2011).

As noted, the success of choice depends on market mechanisms. Hence, (quasi)-markets are created or stimulated in order to facilitate choice. Citizens are supposed to have power in such a market when they are able to shift their expenditure between suppliers as they choose (West, 1998). As a result, a better match between supply and demand and, subsequently, improved public service performance emerges. Empirical evidence indeed suggests that user choice has been found to be positively associated with greater public service performance. For instance, Walker and Boyne (2006: p. 387), analysing the impact of the U.K. Labour government's program, showed that "[…] user choice has a significant positive effect on internal perceptions of service responsiveness, outputs, and outcomes". Savas (1977), using a case of increased competition in refuse collection services in the city of Minneapolis, showed that competition increased productivity and resulted in a more cost-effective service delivery for citizens. Furthermore, evidence from the health care sector showed that mortality rates fell in more competitive markets (Cooper et al., 2011).

However, markets do not always operate in the way they are ideally supposed to. One important characteristic is that citizens do not have options to choose from as organisations have considerable market power. The most extreme and visible option here is a monopoly, where one organisation provides all services. However, there are also more subtle forms of market power. For instance, little choice will be available if the form of care is under-supplied, such as in Dutch marketized child care (Kremer and Tonkens, 2006). As almost all organisations have waiting lists, parents do not have real choice options. Related to this, local monopolies may exist. These can also be created, for instance when hospitals merge, giving citizens less options and results in higher prices (Le Grand, 2007: p. 116).

Increasing inequality among citizens

A crucial concern of the opponents of the choice-movement is that introducing choice into public service delivery is increasing inequalities in service provision (Butler, 1993). However, on the supply side, evidence does not suggest that competition between service providers increases inequality. For instance, Cookson et al. (2010) analysed hospital competition in the U.K. and concluded that the behaviour of hospitals and doctors was not increasing socioeconomic health care inequality. Lacireno-Paquet et al. (2002) analysed school choice in the U.S. and showed that market-oriented schools were, contrary to expectation, not focusing exclusively on an elite clientele, although they did serve high need populations somewhat less. However, evidence from the demand side points in the direction of making informed choices strongly depends on socioeconomic status and service users' experiences in making choices. In the end, this may indeed lead to increased inequality.

In the education sector, it has been found that school choice has social segregation effects, leading to children from lower socioeconomic backgrounds being worse off (Hsieh and Urquiola, 2003; Musset, 2012; Howell, 2004). Within the healthcare sector evidence is more ambiguous (Dixon and Le Grand, 2006; Dixon et al., 2007). In the area of utilities it has been noted that potential vulnerable service users are not only less likely to switch their provider (Jilke, 2013), but are also less satisfied with the services they receive (Clifton et al., 2011a; Wilson and Price, 2010).

Why do lower socioeconomic classes have more problems in exercising choice? First of all, we must state that lower socioeconomic status groups face a number of related constraints when exercising choice making it hard to pin-point one particular factor (Hsieh and Urquiola, 2003). However, considering choice options, the notion of switching costs can partly explain the differences (Arksey and Glendinning, 2007; Lent and Arend, 2004). Switching costs are the monetary and non-monetary expenses that a citizen has to pay when he or she changes providers (Burnham, Frels and Mahajan, 2003). The higher the switching costs, the more difficult it becomes for citizens to exercise choice.

Important switching costs in the choice debate are procedural switching costs. Procedural switching costs consist of economic risk, evaluation, learning, and setup costs, and primarily involve the expenditure of time and effort (Burnham, Frels and Mahajan, 2003). These procedural switching costs are higher for less educated, older and the mentally handicapped; thereby increasing social inequalities (Arksey and Glendinning, 2007; Lent and Arend, 2004). Meinow et al., (2011) found that older people do not have the necessary capacity to collect and evaluate information for making choice decisions. Lako and Rosenau (2009) found that most patients do

not independently choose a hospital based on available performance information. Rather, they rely on other sources, such as recommendations from their general practioner, hospital reputation in general, or the distance from their home to the hospital. They base their information on so-called information networks. However, such networks vary in accordance to its members' socioeconomic class, with lower socioeconomic groups having poorer networks. Furthermore, evidence in the US education sector points to introducing choice having a positive effect on the nature of information networks; however, they were associated with higher levels of class stratification and racial segregation (Schneider et al., 1997).

Related to this are risk-aversive switching costs. It is evident from various studies that people tend to stick with the default, the service provider they are already using (Wilson and Price, 2010; Jilke, 2013). The status-quo is a safe haven, a so-called satisficing option. Higher educated service users are more likely to exhibit greater risk seeking behaviour, while their lower educated counterparts are risk avoidant (Dohmen et al., 2008) and thus tend to stick with their current provider.

Problems with using information: Bounded rationality and gaming

Another, related pitfall concerns the way in which performance information is presented. A major problem of using performance information is the bounded rationality of clients. Service users are bounded rational, even if the full information was available to them (Simon, 1955). Parents, patients, clients, or service users employ the same heuristics and mental shortcuts when making choice decisions as they do during their daily process of decision making. This includes information overload, simplification heuristics, risk aversion, or status-quo bias, among many others (see for an overview Kahneman, Slovic and Tversky, 1982). Too much information may confuse service users and result in oversimplification, using other, seemingly irrational, criteria than quality or price to determine their decisions, for example reflecting the choices within their social networks. Service users rather then rely on hear-say than on league table figures (Marshall et al., 2000).

One example of bounded rationality is that providing greater opportunities to choose from does not necessarily lead to more active choice behaviour. Studies in the area of applied psychology indicate that the effects of increasing ones' choices on buying decisions follows an inverted U-shape (Shah and Wolford, 2007). In other words, increasing the number of alternatives has first a positive effect on buying decisions in general, but too much choice may overwhelm service users resulting in choice avoidance and dissatisfaction (Schwartz, 2005). In a famous experiment, Iyengar and Lepper (2000) showed that people are more likely to purchase gourmet jams when offered a limited array of 6 choices rather than a more extensive array

of 24 choices. Moreover, people actually reported greater subsequent satisfaction with their selections and wrote more positive reviews when their original set of options had been limited to 6. Hence, increasing opportunities can become 'too much of a good thing'. While one may argue that the number of alternatives is not as pronounced in public service provision as it is the case for private goods, first evidence is available from the field of liberalised infrastructure services, which suggests just the opposite (Jilke, 2013). The application of these results in the area of healthcare, social service provision or education should be examined.

Next to bounded rationality, a second problem with using performance information is suppliers' gaming behaviour. Service providers sometimes engage in gaming with figures (De Bruijn, 2007). Hood (2006) shows that suppliers in British public services indeed played extensively within the targets set by Tony Blair's New Labour Government. For instance, he noted that "*In studies of an eight-minute response time target for ambulances dealing with Category A calls (life-threatening emergencies), there were large and unexplained variations in the proportion of calls logged as Category A, and ambiguity over when the clock started*" (2006, p. 517). Hence, it seems that suppliers manipulate (play with) performance information making it less reliable.

Worsening work conditions

The last pitfall we discuss, concerning worsening work conditions, seems to attract far less attention than the previous three pitfalls (Ungerson, 2004; Kremer and Tonkens, 2006; Ungerson, 1997). According to choice protagonists, power should shift from organisations and employees towards clients. Studying the introduction of choice in care settings, Ungerson (1997: p. 46) notes "The evidence is overwhelming that disabled people have in the past been demeaned, discriminated against, abused and ignored by precisely those people funded by the state who were and are supposed to respond to their needs". However, this shift in power could have severe consequences for employees.

First, the introduction of (quasi)-markets needed for choice to operate successfully could lead to a 'grey' labour market, marginalising employees and locking them into low-paid and transient employment (Ungerson, 1997). Related to this, Knijn and Verhagen (2007) showed that introducing client choice via personal care budgets leads to increased managerial demands and work pressure for employees. This is especially true for employees who were, also prior to the introduction of choice, rather powerless. In this respect, Ungerson (1997) discusses personal assistants in care settings in the U.K. Personal assistants are carers employed on a short or permanent basis by the patient. They sometimes live in the same household as the

patient in order to be readily available. Ungerson notes that with the introduction of choice, these personal assistants work in a grey labour market, which is unregulated and underprotected. In her view, this is likely to add to poverty rather than decrease it. This is especially problematic given that female, old, or immigrant workers, who are already in a less favourable position in general, are highly likely to take up these jobs. Another example is the regulation of the postal market. Evidence from Germany and Austria shows that increased competition has led to lower wages and less job security for postmen (Schulten, Brandt and Hermann, 2008).

Additionally, the introduction of choice can negatively affect the professionalization of employees (Knijn and Verhagen, 2007). Firstly, it challenges the autonomy of employees. This is especially relevant for groups who are not regarded as traditional professions. For these semi-professionals, it becomes more difficult to be critical towards a citizen, when that citizen directly pays you. Secondly, there is a threat that there will be less development of professional knowledge. In the Netherlands, employees working using Personal Care Budgets note that employees are worried about their professional development. They miss direct contact with peers and complain about the lack of space for developing their knowledge and education (Kremer and Tonkens, 2006).

5. Conclusion

The aim of this chapter was to provide an overview of the background, facilitators and pitfalls of choice. As has been shown by various empirical studies discussed in this chapter, introducing choice in public services can have benefits, such as increased public sector performance. Furthermore, choice can empower citizens, as the example of personal care budgets shows. However, it has been shown that choice also has pitfalls, such as worsening work conditions for employees, and problems with appropriately using performance data. More importantly, choice can also be unequally divided; clients with lower socioeconomic status being worse off. Such a situation is especially problematic in public services. In this vein, Hood (1991: p. 3-19) talks about the importance of so-called 'theta' values in public services, such as fairness, honesty and mutuality. These can be under attack when choice is introduced.

Based on the analyses, we argue that choice is not 'a good in itself' (cf. Giddens, 2003). In our view, choice is a means to an end. Further research could reflect on the introduction of choice, analysing the advantages and disadvantages in particular contexts. Furthermore, much can be learned from combining evidence of various

studies in different sector, using systematic reviews or meta-analytical techniques. Scholars could conduct meta-analyses on important issues in the choice debate which were discussed in this chapter, such as the relationship between a) choice, market power and performance; b) choice and inequality among citizens; c) choice and the use of performance information; and d) choice and work conditions.

Based on the results presented in this chapter, we would advise policymakers to make informed decisions when introducing choice in public services. We are not saying that policymakers should never introduce choice, as it can have substantial advantages. Furthermore, policymakers could take measures to avoid falling in a pitfall. If policy makers were to do this systematically, we believe that it would substantially enhance the effectiveness and legitimacy of introducing choice in public services. As such, we promote a continuous review of the effects of choice throughout its introduction. In this way, choice can be introduced in some public services, and many can reap its benefits.

CHAPTER 3

Public Management Reform and Responses to Decline in Public Services: An Experimental Evaluation of Choice-Overload

This chapter has also appeared as Jilke, Van Ryzin and Van de Walle (2014).

1. Introduction

Policy makers and academic proponents of introducing competition and choice into public service delivery have repeatedly claimed that overcoming state monopoly-led provision of public services would increase the efficiency of public service delivery and result in a boost in citizens' welfare (Ostrom and Ostrom, 1971; Savas, 1987). It has been argued that this would be achieved through shifting the autonomy for decision-making (in terms of provider choice) from the state to the citizen by creating (quasi-)markets for public services and allowing service providers to compete for customers. Citizens would then send market signals to suppliers by complaining to or switching their providers. Service providers, in turn, would respond accordingly by adjusting the value-for-money of their services in order to keep service users as well as attracting new ones. As a result, the long-run equilibrium would be achieved between citizens' demands and preferences and the price and quality of the offered services.

Such a neo-classical perspective on public service delivery under competition tends to assume that increasing the number of service providers to choose from will result in an optimal allocation of available resources. Or in other words, public services would become cheaper and better. This rests on the assumption that citizens choose from a set of service providers the one that best matches their needs and demands (Stone, 2005). However, can there be too much choice to do so? In a well-known study, Iyengar and Lepper (2000) conducted a field experiment where they randomly assigned customers of a US grocery store to taste among a set of six (limited choice), or twenty-four (extensive choice) different gourmet jams. Subsequently, the authors found that despite the perceived higher attractiveness of the larger sample of jams, people who were exposed to the extensive-choice condition were clearly less likely to purchase one of the jams. Or in other words, increasing choice reduced people's likelihood of making a decision. These findings stand in stark contrast to basic assumptions put forward by standard psychological theories of human motivation and economic theories of rational decision-making, that is "[…] that having more, rather than fewer, choices is necessarily more desirable and intrinsically motivating" (Iyengar and Lepper, 2000: p. 997).

Although the choice-overload effect has been studied many times in various private-sector contexts (for an overview see Scheibehenne, Greifeneder and Todd, 2010; but see also Cherney, Böckenholt and Goodman, 2010; Gonzales, 2013), commentators have questioned whether it matters in the case of public service markets where only a limited amount of service providers compete for customers (Dowding and John, 2009; Le Grand, 2007). But there has been push in many countries to

liberalize the provision of core public services, such as education and health care, in order to create more choice and competition. A prime example is the provision of electricity, which used to be delivered by state-owned or state-regulated monopolies that gave residents of a city or region essentially no choice in providers. But nowadays the electricity markets have been liberalized and/or de-regulated to a great extent (Conway and Nicoletti, 2006). This means that in many markets there are now multiple public and private service providers that compete for electricity customers. For example, in the State of New York people have on average 41 electricity providers available to choose from[11], clearly a situation in which choice-overload could well be a relevant factor. But while the possible adverse effects of choice-overload for public service provision have been discussed repeatedly (Dowding and John, 2009; Lipsey, 2007; Macaulay and Wilson, 2008; Schwartz, 2004; Tummers, Jilke and Van de Walle, 2013; Jos and Tompkins, 2009; Wilson and Price, 2010), an empirical evaluation of the unintended negative consequences of increasing provider choice is missing in the public management literature.

In this study, we extend and test the theory of choice-overload by investigating whether or not increasing the number of providers of public services in the US electricity sector has adverse effects on peoples' motivation to switch their provider after a service failure. To do so, we use a survey experiment (N=1,154) were we randomly vary the number of service providers in a hypothetical service failure scenario. Results show that increasing provider choice reduces people's likelihood of switching away from a poor performing provider by about 10 percentage points. These findings also hold when replicating the experiment with an independent online sample (N=541). Thus our results indicate that increasing provider choice in public service markets causally influences peoples' motivation to switch away from poor performing public services. In turn, this may lead to a situation where citizens are being locked-in to a suboptimal provider simply due to an overload of choices.

2. Choice-Overload and Public Management Reform: A Theory for Citizen Responses

Empirical studies on the detrimental effects of too much choice have spread considerably since Iyengar and Lepper's (2000) seminal jam-study in which they found that offering people too much choice reduces their motivation to choose.

[11] We counted the number of service providers for each of the 62 cities within the state of New York via www.newyorkpowertochoose.com. The numbers are for March, 2014. If the number of providers varied across city zip-codes, we took the number available within the respective city centre. When there were multiple grids available for citizens to choose from, the mean value of providers across grids was used.

Since then, various studies in social psychology and marketing have corroborated a choice-overload effect in different contexts, ranging from simple consumption decisions for items such as chocolates (Cherney, 2003), pens (Shah and Wolford, 2007) or gift boxes (Reutskaja and Hogarth, 2009), to more complex decisions like choosing music players that differ on many attributes (Greifeneder, Scheibehenne and Kleber, 2010), volunteering with a charitable organisation (Caroll, White and Pahl, 2011), or enrolling in 401(k) pension plans (Iyengar, Huberman and Jiang, 2004). Studies have also shown that having too many choices not only undermines people's motivation to choose, but also negatively impacts their subsequent satisfaction with the option they have chosen (Diehl and Poynor, 2010; Greifeneder, Scheibehenne and Kleber, 2010; Haynes, 2009; Reutskaja and Hogarth, 2009), including disappointment and regret (Schwartz, 2000). Proponents of choice-overload argue that these adverse outcomes can be explained by three basic factors: information overload, unclear preferences and negative emotions (for an overview see Botti and Iyengar, 2006).

Research on information overload suggests that individuals have limited capabilities to encode information, and when those limits are reached people tend to become uncertain (Chen, Shang and Kao, 2009; Lee and Lee, 2004; Miller, 1956; Timmermans, 1993). Therefore, as the amount of information to be processed grows, decision-making becomes poorer and the motivation or ability to make a decision diminishes. Moreover, work in the field of social psychology has shown that people do not hold stable and clearly ordered preferences ready at their disposal when faced with a choice; rather, people's preferences are fluid and heavily context dependent (Feldman and Lynch, 1988; Kahn and Baron, 1995; Payne, Bettman and Johnson, 1993; see also Botti and Iyengar, 2006). Given this fluidity and contextual nature of people's preferences, having to choose among a large array of alternatives can produce cognitive conflicts and overload, which can in turn result in negative emotions and stress (Botti and Iyengar, 2006). In particular, choosing in a context of too many options often means disregarding potentially attractive alternatives, and regretting forgone choices has been found to be associated with choice deferment (Beattie et al., 1994; Simonson, 1992). As a result people often avoid choosing altogether.

Closely related to this stream of research is Simon's (1955) concept of bounded rationality. Here it is assumed that the decision-making of individuals is bounded by individual-level restrictions, including uncertainty, cognitive constraints in processing information, and information overload. Because of such personal limitations, people may end-up making poorer (non-optimal) decisions, including sticking with their default. In other words, according to the concept of bounded

rationality, individual's risk of making a poor decision increases when the number of options to choose from goes up. Since people are in general risk-averse, and since choosing among many options is often seen as a risk of making the wrong choice, they tend to avoid making a decision at all. In such situations, people will tend to stick with the default, a tendency Samuelson and Zeckerhauser (1988) refer to as 'status-quo bias' (see also Thaler and Sunstein, 2008). In the context of public service provision, status-quo bias would imply that citizens will become more likely to stay with their current service provider when faced with too many alternatives in the market.

However, as suggested, such negative effects of too much choice would seem to be context dependent. In their meta-analysis of choice-overload effects in consumer research, Scheibehenne and colleagues (2010) find that despite great variance between studies, there was no support for a single main effect of choice-overload (but also see Cherney, Böckenholt and Goodman, 2010; Gonzales, 2013). They conclude that the great variance between studies could be a result of the context dependent nature of choice-overload. Yet, in their meta-analysis, they were not able to identify any specific preconditions that must be met for choice-overload to occur. Thus, in line with the critiques of their meta-analysis (Cherney, Böckenholt and Goodman, 2010), Scheibehenne and collaborators (2010) call for further research to identify such conditions. Indeed, most of the studies examined in their meta-analysis drew upon simple consumption decisions of private goods with relatively simplistic attributes and limited daily importance to participants (such as jams or pens). Our study differs from these existing works in two important ways. First, we study people's responses to service decline instead of a simple consumption decision based on various product attributes; and second, we focus on a core public service (electrical power) that is fundamental to modern life and of great daily importance for citizens. While much of the research on choice-overload has indeed concentrated on simple consumption decisions for private goods such as jams, pens, chocolate, or music players, we do not know whether the choice-overload hypothesis holds true for citizens' responses to poorly performing public services. This is an important question since citizens cannot simply withdraw from using public services such as electricity, but need to switch providers if they want to leave poor performing service organisations. Such a choice decision is arguably much more meaningful than deciding whether or not to buy a gourmet jam or a pen, although both types of decisions may be susceptible to choice-overload.

Theories of citizen responses to decline in public service performance suggest that as a result of their dissatisfaction with public services, citizens frequently switch between public service providers (Dowding and John, 2012; Lyons and DeHoog,

1992). The key driver here seems to be their dissatisfaction with the state of affairs of a certain public service. This assertion rests on Hirschman's (1970) classical distinction between exit and voice as response to organisational failure, meaning people can either voice their dissatisfaction, hoping that things improve, or they can leave the respective organisation or service provider. The likelihood of exit and voice, in turn, is moderated by people's loyalty to the organisation in question. While voice can be either collective – such as participating in a demonstration, or electoral voting – (see for example James and Moseley, 2014; Boyne et al., 2009) or individual – such as filing a personal complaint – (see for example Jilke and Van de Walle, 2013), exit means that people either stop using the service in question or switch to another (public or private) service provider (for an overview see Dowding and John, 2012).[12] In this study, we focus on exit in the form of switching providers as a response to a decline in public service performance.

Public management reforms over the past decades have often aimed at improving citizens' opportunities to choose among multiple providers of public services. Indeed, a core element of the New Public Management inspired initiatives was to move away from state-led provision of public services to opening up public services markets for (quasi-)competition (Pollitt and Bouckaert, 2011; Osborne and Gaebler, 1993; Barzeley, 2002). This holds especially true for the public utility sector (Bognetti and Obermann, 2008). While in the 1970s public or heavily-regulated private monopolies were omnipresent, nowadays effective competition is apparent in many public service sectors across Europe and North-America (Conway and Nicoletti, 2006). Thus, these reforms have substantially changed the way public services are delivered today. Indeed, a central aim of liberalising public service provision was to insert choice into the provision of public services and foster competition among providers. Le Grand (2007) argues that choice can lead to greater responsiveness to service users' needs and demands, but only if it meets two conditions: 1) competition must be real (there must be real competitors), and 2) there must be a real choice among alternatives. But even if choice and competition may not lead to desirable outcomes, Le Grand (2007) makes the case that choice has intrinsic value on its own. For citizens, however, simply increasing the number of alternatives does not necessarily leave them with more valued choice, as Dowding and John (2009) argue. For example, if parents could choose among a public and private school for their kids, adding another similar public school would not increase their options as much as, for example, adding a faith-based school. The choice-set of the three very different types of schools would be valued higher than a choice-set of three very

[12] A third alternative would be a physical relocation (Dowding and John, 2012). Its most famous subset is the so-called Tiebout-exit. However, one has to note that in its original form, a Tiebout-exit is a response to relatively high municipal taxes (Tiebout, 1956).

similar schools. Thus parents' indirect utility would be most likely higher for the first option, albeit having the same number of alternatives. But while "[...] increasing the number of alternatives ('hard choice') does not entail increasing choice in any valued sense, it may induce costs" (Dowding and John, 2009: p. 228), including psychological costs.

In line with Hirschman's (1970) theory of responses to poor performing organisations, we argue that people respond to a decline in public service performance by either expressing their discontent (voice), or leaving the public service provider in question (exit). However, citizens' likelihood of exiting is decreased when there are too many alternatives to choose from. While people can rather easily navigate through a set of two or three alternatives, the growing overload of information that comes with many alternatives produces cognitive conflicts, including stress, and makes citizens uncertain about picking the 'right' option. Anticipating the risk of potentially being worse-off after switching (loss aversion), individuals stick with the service providers they are currently using (status-quo bias), even if they are dissatisfied with their services. This assertion is in line with key tenets of choice-overload, but also provides a valuable extension of the theory by not only considering simple consumption decisions, but applying it to models of citizens' responses to poor performing public services. Therefore, the central hypothesis we test in this study is as follows: All things being equal, citizens who experience severe dissatisfaction with a given service will be less likely to switch away from their current service provider when faced with many alternative providers, compared to people who are equally dissatisfied but have a more limited set of providers to choose from.

3. Experimental Design and Participants

To investigate the choice-overload hypothesis in the context of public service delivery, we designed a discrete choice experiment based on a hypothetical service failure scenario (see also Maute and Forrester Jr., 1993), which was embedded in an online survey. The particular strengths of survey experiments is that they combine both, the internal validity of classical experiments and the external validity of population surveys (Mutz, 2011). This allows us to make a firm cause-effect assessment of choice-overload across a very diverse subject pool. In this context, we examine our theoretical predictions in the context of enhanced deregulation and competition in the provision of electrical power in the US. Historically, the US electricity market was dominated for much of the 20th century by monopolistic municipal utilities, power cooperatives, or privately held companies regulated by public utility commissions. Following liberalisation trends that began in countries

like the UK and Chile, coupled with advances in smart-grid technology, the US electricity market over least few decades has experienced deregulation and a proliferation of the number of electricity providers available in many states and metropolitan areas. As mentioned earlier, New York State now includes over 40 electricity providers on average, and other states with deregulation and significant retail choice include Texas, Pennsylvania, Ohio, Illinois, New Jersey, and Michigan (US Energy and Information Administration, 2010). Thus, choosing an electricity provider is a necessary and salient task faced by a great many people in the US and makes this a relevant public service context for studying the choice-overload hypothesis.

Sample: Amazon's Mechanical Turk

For our study, we used Amazon's Mechanical Turk (MTurk)[13] to recruit experimental subjects. Being an online labour market platform where respondents are paid for small tasks, including survey participation, scholars have increasingly relied on the MTurk for doing experimental and/or survey research. Various studies have demonstrated the quality of data obtained via MTurk (e.g. Amir, Rand and Gal, 2012; Berinsky, Huber and Lenz, 2012; Crump, McDonnell and Gureckis, 2013; Goodman, Cryder and Cheema, 2013; Horton, Rand and Zeckhauser, 2011; Paolacci, Chandler and Ipeirotis, 2010). Nevertheless, the MTurk pool of online workers is clearly not a random sample of the population and hence not statistically representative. Nevertheless, the MTurk population is very diverse in terms of demographic population parameters when compared to other non-random samples that are regularly used for experimental studies, such as college student samples, or even standard internet panels (cf. Buhrmester, Kwang and Gosling, 2011). Moreover, scholars have used samples from the MTurk population to replicate studies that were produced using traditional random samples and have found no substantial differences in the results obtained. For example, Berinsky, Huber and Lenz (2012) replicated two key experiments in political science using the MTurk. The first one comes from a randomly recruited and nationally representative face-to-face survey in the US, focusing on the topic of question wording in the area of welfare spending. The second was an experiment on risk-orientation and framing effects that used a Knowledge Networks probability sample. For both MTurk experiments, the replications yielded highly similar results to those obtained from the original experiments using a probability sample. But experimental results from MTurk are also consistent with findings produced in behavioural laboratories, which are commonly regarded as the gold-standard in terms of experiments' internal-validity (Kagel and Roth, 1995). Suri and Watts (2011) have replicated the findings from a

[13] For more information see https://www.mturk.com/mturk/welcome.

series of public goods experiments via the MTurk which initially had been conducted in the classroom. They have found no significant differences in obtained results. The same holds true for Horton, Rand and Zeckenhauser's (2011) MTurk replication of a one-shot prisoner's dilemma game that had been conducted previously in a laboratory setting. Berinsky and colleagues (2012) report on a successful replication of Tversky and Kahneman's (1981) classical "Asian disease problem" which was originally conducted using a student sample, but since then replicated in many different settings. In sum, "[…] evidence that Mechanical Turk is a valid means of collecting data is consistent and continues to accumulate" (Mason and Suri, 2012[4]); thus, using MTurk can be regarded as a promising sampling frame for experiments in public administration research involving a general population.

We hosted our survey-experiment through the Qualtrics software and directed subjects to the URL provided in their work-description. Only US-based participants were recruited. One concern when using online recruitment panels is that subjects rush through the online questionnaire without properly reading the provided instructions and questions. Indeed, Goodman, Cryder and Cheema (2013, p. 213) illustrate that participants recruited via MTurk "[...] are less likely to pay attention to experimental materials". Others also warn against so-called bots, computer programs who are programmed to answer survey questions (Mason and Suri, 2012). Therefore, to increase the statistical power and reliability of our dataset, and to reduce Type II error (false negatives), we screened respondents based on two criteria. First, we included an instructional manipulation check[14], as recommended by Oppenheimer, Myvis, and Davidenko (2009) to detect 'satisficers', spammers, or even bots. Those study participants that failed this attention check were excluded from our sample (a total of 30 respondents). Second, we examined the time subjects took to fill out the questionnaire (mean of 5.23 minutes, with a standard deviation of 3.37). Extreme deviations from the average time to complete the questionnaire were regarded as an indication of satisficing behaviour that was not captured by the instructional manipulation check (see also Mason and Suri, 2012; Mutz, 2011). Thus respondents within the highest and lowest 1% percentile (less than 1.77 and more than 17.92 minutes respectively) in terms of total time till survey completion were excluded (23 subjects in total). Another commonly raised concern about using MTurk samples is that respondents log-in to the online platform with multiple accounts and participate in the very same experiment more than once (Chandler, Mueller and Paolacci, 2014; Horton, Rand and Zeckhauser, 2011). This obviously violates the assumption that subjects are independent observations and thus poses

[14] More precisely, we presented respondents with the following question: "*To ensure participants read the questions, please select "Very Satisfied" on the scale (first option)*". Then they were presented with a horizontally ordered five point Likert Scale ranging from "Very Satisfied" to "Very Dissatisfied".

a threat to the internal validity of the experiment. Hence, we checked whether the subjects' internet protocol (IP) addresses overlapped, as proposed by Horton, Rand and Zeckenhauser (2011). When this was the case, subjects were excluded from our analysis (11 respondents in total). After applying these screenings, we were left with a total number of 1,154 study participants.

Table 3.1 presents the characteristics of our sample of respondents from the MTurk with regard to gender, age, income, race and place of residence. Respondents are more likely to be white, male and younger, but represent a range of incomes and places of residence. Although not representative, the sample is nationwide in scope and fairly diverse.

Experimental procedure

To assess the choice-overload hypothesis in the context of liberalised public services, we have conducted a scenario-based survey experiment based on a 2x2 factorial, between-subjects design as depicted in Figure 3.1. Dissatisfaction with the service and the amount of choice available were each experimentally manipulated in the following manner. First, the degree of service failure was experimentally

Table 3.1: Study participants' socio-demographic characteristics MTurk (N=1,154).

	Frequency	Percent
Gender		
Male	734	63.6
Female	420	36.4
Age		
18-34 years old	910	78.9
35-64 years old	235	20.4
65 years or older	9	0.8
Income		
25.000 USD or less	342	26.64
25,000 – 75,000 USD	570	49.39
75,000 USD or more	242	20.97
Place of residence		
Urban	220	40.4
Semi-urban	211	38.8
Rural	114	20.9
Race		
Black or Afro-American	50	4.33
Hispanic or Latino	75	6.50
White, Non-Hispanic	899	77.90
Other	130	11.27

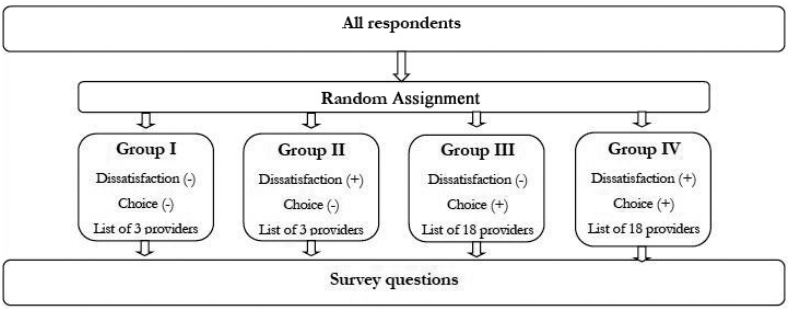

Figure 3.1: Experimental design.

manipulated, ranging from a small to a severe decline in service performance (1st treatment). This was expected to induce subjects to be dissatisfied with the hypothetical service provider. Second, respondents were randomly assigned to a public service market with either a high or low degree of provider choice (2nd treatment). The information provided mimics information that is typically provided on provider comparison websites. The respective scenario description was as follows:

Say you are a resident of Middletown and you receive your home electricity services from ABC-utility, which is owned and operated by the municipality. In recent years, the local electricity market was opened up for competition. There exist [three/eighteen] electricity providers.

Recently, ABC-utility mistakenly over-charged you and deducted too much from your bank account. After notifying them, it took ABC-utility [two/90] days to refund you the money. The customer service representative you talked with at ABC-utility was [friendly and helpful/not that friendly or helpful], and the utility [later/never] sent a letter of apology for the mistake. Surveys indicate that people are in general [fairly/not that] satisfied with ABC-utility.

After being presented with this information, subjects were presented a list of hypothetical service providers, including their current provider (ABC-utility), and a list of alternatives (2 alternatives in the low choice group, and 17 alternatives in the high choice group). These providers differed on various attributes such as the number of customer complaints per year, price, minimum term, and cancellation fee. This was done to show subjects that they have a choice among a diverse set of providers, while avoiding asymmetrically dominated alternatives (see also Huber, Payne and Puto, 1982). The number of complaints per year was regarded as a proxy for service quality. The incumbent (ABC-utility) was assigned the highest or lowest value (230, or 90 complaints per year respectively), in accordance with the satisfac-

tion treatment subjects were assigned to. In the low choice condition, we randomly assigned one alternative the highest/lowest value respectively, and the other the mean value of 160. For the high choice condition values were determined randomly with the range of 90 to 230 complaints per year. Next, we wanted to control for the economic-effects that respondents simply choose the cheapest offer. Therefore, the actual prices were varied between 0.0111 and 0.0127 cents per kwh. The incumbent was assigned the mean value of 0.0119, while all other providers were randomly assigned a value in between 0.0111 and 0.0127. Lastly, we included two additional attributes to signal low switching barriers: the minimum term for each supplier, and the cancellation fee, which are typically available on electricity provider comparison websites[15]. Here we assigned our incumbent a minimum term of 1 month and no cancellation fee, while all other providers were randomly assigned on these attributes[16].

After being presented with this information, all respondents were asked whether they would stay with their current provider, or choose one of the others (which they had to name). This resulted in a discrete choice outcome variable for our subsequent analysis[17]. Respondents were also asked to rate their satisfaction with ABC-utility and the perceived amount of choice available to them as a manipulation check.

Before fielding the actual experiment, we conducted a pre-test via MTurk (N=208) to determine the actual number of providers that respondents perceive as many or only a few choices. Here we varied the number of service providers (eight, thirteen, eighteen and twenty-three) based on real-life information. In the State of New York, for example, the mean number of electricity providers within cities is 41 (with a standard deviation of 11.4, ranging from 1 provider in Long Island to 68 in New York City[18]). Our pre-test clearly identified eighteen providers as the number that participants start to consider as a lot of choice when compared to the baseline of three providers. We then pre-tested the final questionnaire among a small sample of MTurk respondents (N=88) which resulted in minor changes in the questionnaire and the detection of some typos. Respondents from the MTurk pre-tests were not included in the experiment that followed.

[15] See for example http://www.newyorkpowertochoose.com.

[16] Minimum term varied between "No", "1 Month", "2 Months", "3 Months", "4 Months" and "5 Months", and cancellation fee between "No", "$10 for each remaining month" "$20 for each remaining month", and "$50".

[17] Study participants were asked "*Below you find a list of all [three/eighteen] electricity providers that operate in Middletown. Remembering the described scenario, what would you do? (Please scroll down)*". Respondents could either indicate that they would stay with ABC-utility, or switch to one of the mentioned service providers (should indicate the provider they want to switch to).

[18] As of March, 2014.

4. Results

Before turning to the main results, we first present evidence of the effectiveness of the dissatisfaction and choice manipulations. As intended, respondents in the severe service failure condition reported significantly higher levels of dissatisfaction than those assigned to the mild service failure scenario ($F=29.84$, $p<0.05$)[19]. Similarly, subjects that were randomly assigned to the low choice condition reported having significantly less choice available in Middletown's electricity market when compared to those participants that were in the high choice condition ($F=6.10$, $p<0.05$)[20]. Thus both manipulations seemed to have worked as intended.

To test the independent and combined effects of the experimental manipulations, we estimated a binary logit regression model with subject's discrete choice decision (stay with current provider versus switching) as dependent variable (Table 3.2). The first model serves as a baseline model. The second model displays the main effects of both treatment variables and tests the choice-overload hypothesis by interacting them with each other. This way we can assess the effect of service failure on respondents' stated switching behaviour contingent on the amount of choice that was made available to them. Both treatment variables have been effect coded in order to avoid misinterpretations of the main effects of the treatments in the interaction model. The corresponding coefficients are conventional logit coefficients and need to be interpreted as representing the difference between the mean of the group that received the treatment and the grand mean of all groups. Or in other words, they display the effects of being in one of the four experimental condition versus all others.

Table 3.2: Experimental results (MTurk; N=1,154).

	Baseline	Choice-overload
Choice	0.432* (0.235)	0.136 (0.268)
Dissatisfaction	5.393** (0.257)	5.485** (0.268)
Choice X Dissatisfaction	–	−1.536** (0.535)
−2 Log likelihood	571.080	562.373
Δ −2 Log likelihood	–	8.71**
Pseudo R-squared	0.64	0.65

*Standard errors in parenthesis; ** p<0.05, * p<0.1 (two-tailed tests, except for the interaction term where a one-tailed test was used because it represents a directional hypothesis)*

[19] Subjects have been asked "*Thinking about ABC-utility, how much you think you would be satisfied with their services, overall*?" and then were given a slider ranging from 0 to 100 to express their satisfaction.

[20] Respondents have been asked "*And how much choice one has in terms of the electricity providers in Middletown, overall*" and then were given a slider ranging from 0 to 100 to show the amount of choice they think one has in Middletown.

The results from our baseline model show that both treatments have an effect on subjects' choice decisions. Being in the dissatisfaction treatment group significantly increased respondents' probability of switching when compared to the grand mean. Moreover, confronting them with a large array of alternative providers (choice treatment) also increased subject's likelihood to switch. In a second step, we examine the combined effect of both experimental manipulations, and thereby whether experimental participants have experienced choice-overload. First, adding an interaction term between both treatment dummies significantly improved the model's fit to the data: the −2 log likelihood significantly decreased by 8.7. Here, the main-effect for the dissatisfaction treatment stays essentially the same when compared to the baseline model, while the main effect of increasing choice decreases and becomes statistically insignificant. Moreover, the interaction term between both treatments turns statistically significant and exhibits a negative effect direction. This means that those respondents that have been manipulated to experience dissatisfaction with their current provider and, moreover, been exposed to the high choice condition were less likely to abandon their default provider in the experiment. This lends support to the choice-overload hypothesis.

Figure 3.2 illustrates the magnitude of this choice-overload effect graphically – here we have used dummy coding of the treatment variables to identify the conditional effects of both treatments. It shows the marginal effects of being dissatisfied on switching, contingent on the amount of choice that has been made available. The

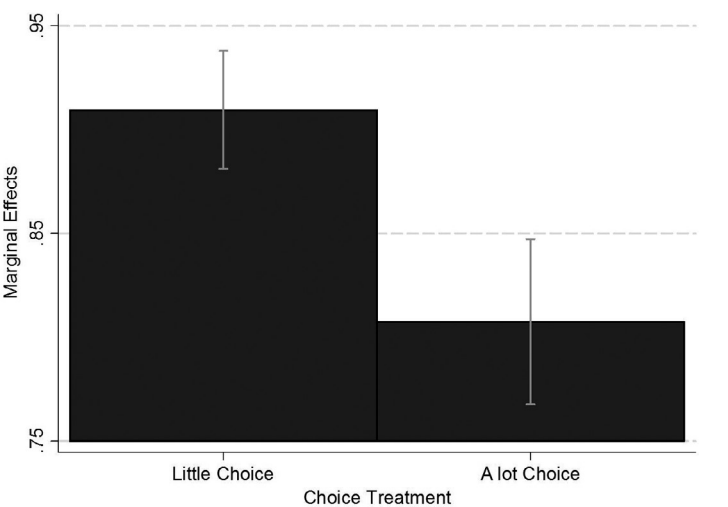

Figure 3.2: Marginal effects of dissatisfaction on switching contingent the amount of choice available (MTurk; 90% confidence intervals).

marginal effect is 0.91 for the low choice condition and 0.81 for the high choice condition. This corresponds to a decreases by 0.1 and is not only statistically significant but also of non-trivial magnitude. In other words, dissatisfied subjects that were confronted with a lot of service providers decreased their likelihood of switching to one of the alternative providers by 10 percentage points.

5. Replication

There have been calls in (experimental) social sciences for an increase in replication studies (see most prominently King, 1995). A very recent large scale replication project in psychology, for example, replicated 13 authoritative studies across 36 different samples with a total of 6,344 participants, and thereby validating a great share of the experimental effects under scrutiny (Klein et al., 2013). Here we aim to validate the experimental results reported in the previous section and, therefore, increase our findings' external validity. Our direct replication used a new sample from an online panel platform of voluntarily recruited participants in the US, called Civic Panel[21]. Participants in Civic Panel are not financially compensated for completing surveys, but rather are included in a post-survey lottery for gift vouchers (10 vouchers of 20 USD, and 1 voucher of 100 USD in total were offered). We used the very same experimental design and Qualtrics questionnaire as reported for the original study. Again, we have performed the same subject screening techniques as have been done for the original experiment, with the only difference that Civic Panel did not provide us the subjects' internet protocol addresses. But since panellists here are not financially compensated, we do not expect that the same respondents answered the survey experiment more than once. Table 3.3 illustrates the characteristics of our sample of 541 participants. Unlike the MTurk sample, the group of participants from the Civic Panel is predominately female and includes fewer young people and more middle-age and older respondents.

Tests of the effectiveness of our experimental manipulations showed that, indeed, both treatments worked as intended.[22] Turning to the main experimental findings, Table 3.4 displays the empirical results from the analysis of respondents recruited via Civic Panel. Indeed, both treatments exhibited an individual effect on people's stated switching behaviour. Being in the experimental dissatisfaction condition significantly increased respondents' likelihood of switching. Similarly, being exposed to a large number of alternatives increased people's probability of switching when compared to the grand mean. Both results are in line with findings from the MTurk

[21] For more information see http://www.civicpanel.org.

[22] Dissatisfaction: $F=6.23$, $p<0.001$; Choice: $F=1.94$, $p<0.001$.

Table 3.3: Study participants' socio-demographic characteristics Civic Panel (N=541).

	Frequency	Percent
Gender		
Male	179	33.1
Female	362	66.9
Age		
18-34 years old	90	16.6
35-64 years old	387	71.5
65 years or older	64	11.8
Income		
25.000 USD or less	89	16.5
25,000 – 75,000 USD	298	55.1
75,000 USD or more	154	28.5
Place of residence		
Urban	216	39.9
Semi-urban	211	39.0
Rural	114	21.1
Race		
Black or Afro-American	46	8.46
Hispanic or Latino	25	4.60
White, Non-Hispanic	444	81.62
Other	29	5.33

sample. Next we include an interaction term between both treatment dummies. The inclusion of such an interaction term increased our models' fit (the chi-square difference between both models is statistically significant at a 10% level). Examining the coefficient for the included term, we find that, indeed, the combined effects of both treatments had a statistically significant effect on subjects' stated choice decision. Here, the standard errors of the interaction term indicate a larger variation than what we found in the MTurk sample. Nevertheless, when inspecting Figure 3.3

Table 3.4: Experimental results (Civic Panel; N=541).

	Baseline	Choice-overload
Choice	0.921** (0.311)	0.935** (0.318)
Dissatisfaction	4.570** (0.311)	4.611** (0.318)
Choice X Dissatisfaction	–	−1.072** (0.637)
−2 Log likelihood	346.594	343.678
Δ −2 Log likelihood	–	2.92*
Pseudo R-squared	0.54	0.55

*Standard errors in parenthesis; ** p<0.05, * p<0.1 (two-tailed tests, except for the interaction term where a one-tailed test was used because it represents a directional hypothesis)*

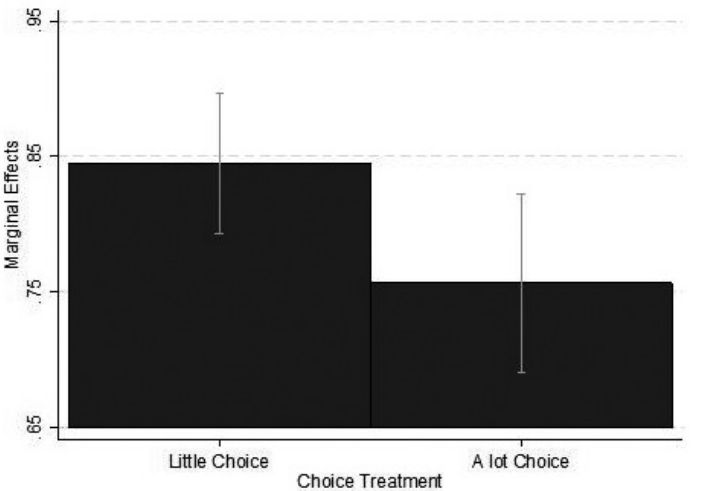

Figure 3.3: Marginal effects of dissatisfaction on switching contingent the amount of choice available (CivicPanel; 90% confidence intervals).

where we depict the marginal effects of dissatisfaction on switching contingent on the amount of choice available, a clear picture emerges. First we can see that the marginal effects of dissatisfaction on switching is 0.85 for the low choice condition. Moving to the high choice condition (0.76), the probability of switching decreases by 0.09. This corresponds closely to the choice-overload effect of 10 percentage points from the MTurk sample. We can therefore conclude that, despite some variation in the identified effects between both samples, we have found a highly similar choice-overload pattern in our replication experiment.

6. Discussion and Implications

This study used a survey experiment and direct replication to empirically examine the choice-overload hypothesis in the context of a key public service, electricity provision, where it has not been tested before. In line with key tenets of choice-overload, we found that dissatisfied individuals who were exposed to many alternative providers decreased their stated preference for switching by 10 percentage points. This effect is not only statistically significant but of potential substantial significance as well, especially in the context of a large population of service users. But before discussing the interpretation and implications of our findings for public services and public management theory, it is important to point some limitations of our methodology and findings.

Although our study used rigorous experimental methods and two diverse national samples of participants, it has important limitations related to the hypothetical nature of the scenario and the discrete nature of the treatments. We manipulated dissatisfaction, in the sense of giving information about a service provider that would lead people to judge it unfavourably, but real failures of an actual public service may induce stronger feelings and thus a greater motivation to switch. As a result, choice-overload may work differently in a real public service market. In addition, we tested decision-making with only two rather extreme conditions, one with only a few providers and the other with a fairly large number of providers. Of course, the effect of choice-overload may well vary along the continuum that lies in between, which probably describes many real-world public service markets in the US and other countries. Thus, it is unclear from these results alone how large the choice-overload effect would be when a city with three providers added, for example, seven more providers. It could be this type of more modest increase in providers results in relatively little choice-overload while greatly expanding meaningful choices and (potentially) satisfaction. It would be useful to have some dose-response experimental designs to probe the point at which choice-overload becomes a significant factor.

With these methodological caveats in mind, we believe our study still has several important implications for public management theory and practice. Previous works in social psychology and marketing have conceptualised choice-overload as a simple consumption decision. Our study expands existing experimental works on choice-overload by applying it to a public service context of service failure through linking the theory of choice-overload with models of citizen's responses to decline in service performance. By doing so it highlights that choice-overload also comes into being when people ought to exit poor performing providers. When too many service providers are present in a public service market, individuals experience choice-overload and thus tend to stick with their current poor performing provider. Choice-overload, in this sense, limits people's ability to respond to organisational failure because of the cognitive biases they face. Yet the ability of citizens to send market signals to poor performing service providers is one of the key assumptions put forward by theories of public service competition. In response, it is assumed that service providers would adjust their services to more closely match citizen's demands and preferences. But given the evidence of a choice-overload effect, the extent to which a long-run equilibrium between citizen's preferences and demands and the offered services will be achieved can be questioned. In addition, the results of this study highlight that increasing provider choice in public service markets can potentially result in consumer inertia; this means that public service users could become locked-in to a suboptimal provider simply due to an overload of choices.

Put together, these findings stand in stark contrast with neo-classical economic thought of individuals acting as rational utility maximisers. Here it needs to be noted that public sector reforms aimed at greater competition and choice have borrowed their inspiration from exactly these theoretical frameworks. Thus our results suggest that theoretical models of competition and choice in public service delivery need to take into account citizen's bounded rationality and associated mental biases.

Our study has shown that reforms of introducing choice and competition into public service delivery are susceptible to overly positive assumptions about citizens' responses to poor service performance. However, the question here should not be whether or not to increase choice and competition in public service markets, but rather how to complement them with appropriate government policies to empower individuals to make those choices that best maximise individual welfare.

CHAPTER 4

Choice and Equality: Are Vulnerable Citizens Worse-Off after Liberalisation Reforms?

This chapter has also appeared as Jilke (2014).

1. Introduction

Although citizens' responses to poorly performing public services have featured on the research agenda for some time (see most prominently Hirschman, 1970; Lyons et al., 1992), recent years have seen an upsurge in interest in studying responses to public services, including citizen satisfaction (Van Ryzin and Charbonneau, 2010; James, 2009), their behaviour when dealing with poor performance (Gofen, 2012; Jilke and Van de Walle, 2013) and the link between these two aspects (Dowding and John, 2011, 2012; Salucci and Bikers, 2011). Following large-scale public service reforms, such as service liberalisation, the pursuit of the 'choice agenda' and the creation of quasi-markets (Clifton and Díaz-Fuentes, 2010; Le Grand, 2007), attention has been drawn to the outcome of these reforms for ordinary citizens (Clifton et al., 2011a, 2012; Florio, 2013; Grosso and Van Ryzin, 2012).

A central aim of these reforms was that citizens, now perceived of as consumers (Clarke et al., 2007), would receive greater value for money through competition among providers. Public service users, in turn, were thought to be able to make well-informed choices and opt for the optimal service provider (European Commission, 2004a). However, it is debated whether all layers of society are equally able to do so. Commentators have claimed that the marketisation of public service delivery and the insertion of greater choice into the public sector might well have fostered a 'two-track' public service where so-called potentially vulnerable service users are less likely to benefit from public service reforms than their relatively 'strong' counterparts (Clifton et al., 2011a; Needham, 2003). Despite these concerns, substantive evidence of negative effects of greater choice on equality in public service provision is lacking.

In France they have a saying "*trop de choix tue le choix*", meaning that too much choice kills the choice (Economist, 2010). While most experiences within the public sector reject the notion that increasing choice necessarily leads to a halt in using the service or an end to switching providers (e.g. Le Grand, 2007), in this paper we go one step further by investigating whether 'too much' choice harms the choices made by lower socio-educational groups – those who are regarded as potentially vulnerable service users. In this study, we examine the cognitive ability element of vulnerability by looking at service users educational attainment. We offer an empirical look at equality in citizens' choice behaviour (switching to another service provider) when it comes to liberalised services of general interest, and particularly in terms of mobile and fixed telephony, in 25 countries of the EU, by asking if potentially vulnerable service users become less likely to switch away from their current service provider once the number of providers increases. Doing

so, this article is structured as follows: the next section introduces liberalisation reforms in services of general interest and studies that have looked at their effects on citizens' attitudes and behaviours. We then address the commonly articulated reproach – that reforms for greater provider choice foster inequalities between service users – and discuss the theoretical and empirical literature regarding reforms in the 'services of general interest' telecommunications sector. Drawing upon the literature on biases in decision-making, we develop our theoretical framework. Subsequently, we introduce our data, measures and methodology, and then test our theoretical framework. Finally, we discuss the findings from statistical tests and extract implications for theory and practice.

2. Liberalising Public Services and Its Effects on Citizens

The European integration process and the creation of a single market fostered the liberalisation of services of general interest and made them subject to greater competition (Héritier, 2001; Prosser, 2005), leading to the creation of liberalised markets for public service provision (Clifton and Díaz-Fuentes, 2010). These markets seek to overcome the market failure situations that typically occur when public services are provided through a monopolistic provider, by establishing a market environment where, ideally, multiple service providers compete for customers (Savas, 1987; Ostrom and Ostrom, 1971). Further, through market signalling, this is expected to create incentives for providers to deliver greater value for money in order to keep existing customers as well as attract new ones. A key attribute in the provision of services of general interest such as water, electricity, or telecommunication services, is that the classical exit option of completely withdrawing from the service in question is often not feasible, too difficult, associated with extremely high costs (see Clifton et al., 2011a), or even may harm citizens' individual welfare. Services of general (economic) interest are, furthermore, regarded as essential public services and "[...] subject to specific public service obligations by virtue of a general interest criterion" (Commission of the European Communities, 2004, citied after Van de Walle, 2008, 7; see also Clifton and Díaz-Fuentes, 2005; 2010). It is because of this general interest character that equality in terms of accessibility and the provision of services of general interest, is regarded as crucial in all EU member countries (Clifton, Comin and Díaz-Fuentes, 2005; Prosser, 2005).

For citizens, changing the delivery and supply arrangements of services formerly provided by public monopolies meant that they were no longer regarded as mere legal subjects, but as vocal and empowered consumers (Aberbach and Christensen, 2005; Clarke et al., 2007). They were put in a position to autonomously make choices

as to which service providers best matched their needs and demands. Experiences in the US telecommunications sector showed that service users were indeed more likely to be better off after switching (Eppling, 2003), while evidence from the UK's electricity market suggests that some service users failed to identify the appropriate supplier for their levels of consumption (Wilson and Waddams-Price, 2010). However, greater provider choice has not always become available within all the liberalised sectors in the EU. The rail transport sector, for example, has remained strongly regulated in most countries, whereas competition and choice is observable in many EU member countries in terms of mobile telecommunications (Conway and Nicoletti, 2006; European Commission, 2010).

The general process of public service liberalisation has been criticised as mainly advantaging the comparatively strong and well-positioned service users, and leaving behind those who are viewed as potentially vulnerable, such as the low educated (Clifton et al., 2011a; also Gottfried, 2001). The literature suggests that comparatively strong and well-educated service users are more prone to take decisions regarding the services they receive that come close to an optimum, when compared to potentially vulnerable service users. This can lead to a service delivery system where potentially vulnerable service users receive least value for money. In terms of services of general interest, numerous observers have shown that various potentially vulnerable service users are indeed least satisfied with the services they receive (Bacchiocchi et al., 2011; Clifton et al., 2011a; Ferrari et al., 2010; Florio, 2013: Fiorio and Florio, 2010; Poggi and Florio, 2009). Moreover, liberalisation reforms have been found to decrease service satisfaction across a whole range of services of general interest (Bacchiocchi et al., 2011; Ferrari et al., 2010; Fiorio and Florio, 2010). However, no clear evidence is available about whether the gap in satisfaction levels between different socio-economic groups increases (or decreases) as liberalisation reforms move on.

Other research on liberalisation reforms has attempted to identify inequality effects on public service users' actual market behaviour or their financial situation within those markets (e.g. Clifton et al., 2011; Jilke and Van de Walle, 2013; Poggi and Florio, 2009). These studies find that inequalities in actual spending, complaints and experiencing financial problems in paying service bills are apparent for numerous groups of potentially vulnerable service users. However, empirically attributing these vulnerability-effects to the liberalisation reforms proves difficult. For example, it may also be possible that potential vulnerable public services users were already in disadvantaged positions prior reforms took place. This clearly justifies further research. Furthermore, within this stream of the literature, there is little evidence available on the extent to which service users, and in particular

those who are regarded as potentially vulnerable, exercise what was argued to be a core element of service liberalisation reforms and greater competition – namely user choice.

3. Equality and Choice Behaviour in Liberalised Public Service Markets

Few studies have examined equality in provider choice by looking at direct and/or indirect effects of socioeconomic aspects on service users' switching behaviours. Ranaganathan et al. (2006) show that young service users are more likely to switch their mobile providers than older users, arguing that this is a reflection of their active market behaviour and high service usage (see also Grzybowski, 2008). Eppling (2002) studied the effects of switching on price discrimination among different users groups. Her results indicated that non-switchers were more likely to have paid higher prices. Her findings further showed that education is positively and income negatively related with switching – that the poor seem to more actively search for better offers. Because of this they also may end up with better offers. Regards education, the author explained her finding by arguing that information is crucial for making choices and finding a better provider, and that the more highly educated service users were more likely to have better access to information. This is so because low search costs are crucial to identify an optimal provider. Moreover, they have greater cognitive abilities to process such information and thus experience greater transaction costs in accessing and processing needed information. Hence, there is considerable evidence that service users who are regarded as potentially vulnerable, such as those who are less well educated, face greater hindrances in making 'optimal' choices because of the increased transaction costs they face in accessing, processing and comparing information. However, one also has to consider the distinct dimensions of vulnerability, income and education tap in: While the relationship between income and choice is rather rational and most likely also related to search opportunity costs, the negative relationship between education and switching stems from structural disadvantages these vulnerable service users face.

The structural disadvantages potential vulnerable services users, such as those with a low level of education, have in choosing among service providers is further reinforced by insights from decision theory. This stream of the literatures suggest that as the amount of information to be processed grows, decision-making becomes poorer and also less likely (Chen et al., 2009; Hwang and Lin, 1999; Lee and Lee, 2004). This is mainly because individuals have limited capacities to deal with information for making decisions, and when those limits are reached, individuals tend

to become confused (Miller, 1956; Timmermanns, 1993). As a result, the likelihood of staying with one's current service provider increases because this represents a safe haven, a so-called 'satisficing' option – a situation which has been more generically described as a 'status-quo bias' (Samuelson and Zeckerhauser, 1988). Related to this idea is the concept of default-effects (Wilson, Garrold and Munro, 2013), where individuals have a tendency to stay with the status-quo even when switching would potentially benefit them. Here it is argued that the reason why people often stick with the default are not only the direct transaction costs such as actual switching costs, but also related to indirect transaction costs such as search costs (see also Wilson, 2012).

Studies in the field of applied psychology indicate that increasing the number of alternatives first results in a positive effect on consumers' choice behaviour, but eventually the effect becomes negative (Botti and Iyengar, 2006; 2004; Iyengar and Lepper, 2000; Schwartz, 2005; Shah and Wolford, 2007), supporting the assumptions linked to information overload. Iyengar and Lepper (2000) were able to identify choice overload in a simple buying environment using a rather large number of alternatives (24), as have later studies (Botti and Iyengar, 2006; 2004). In this regard, Shah and Wolford (2007) highlight the existence of a tipping point (between 10 and 12 alternatives) when there are too many alternatives to choose from, turning choice into too much choice. However, in the case of public service delivery, the number of alternatives or available service providers citizens can choose from is typically smaller – we found a maximum of 13 service providers within mobile, and 10 within fixed, telephony markets.

The fact that individual capabilities in processing information vary among different socio-educational groups, as does the propensity for taking a risk based on a possibly poor decision (Dohmen et al., 2010; Falch and Sangren, 2006; Hjorth and Fosgerau, 2010), means that potentially vulnerable service users are more likely to be risk-averse in terms of their switching decisions. This is mainly so because of their limited capacities in processing and evaluating necessary information, and the higher search costs they face. But choosing among an increasing set of options requires an increase in cognitive effort (Keller and Staelin, 1987). Or in other words, increasing the number of choices affects consumers' search costs to collect and interpret data on a variety of different offers, increasing their indirect transaction costs. Furthermore, in markets with an increasing number of providers to choose from, price discrimination and obfuscation are more prevalent, which further increases the complexity of choosing among a large set of providers. This further increases search costs for service users. We, moreover, argue that citizen-consumers experience differing degrees of search costs, based on their level of

vulnerability. This results in default-effects, and these default-effects increase with their level of vulnerability. If this is the case, then we would assume that, as the number of alternatives grows, the gap between different types of service users will widen. In other words, the difficulty in figuring out the optimal service offer increases as the number of service providers increases. That is, determining the optimal provider becomes more difficult for this particular group of public service users and making a choice then represents a risk to them. As a result, they are more likely to stick with their current provider and tend to become 'locked-in'. This leads to the research question we aim to address in this study: Do potentially vulnerable service users – compared to less vulnerable users – become less likely to switch away from their current service provider once the number of providers increases?

4. The European Telecommunication Sector

In this study, we look into citizens' switching behaviour in a strongly marketised service sector – telecommunications. The European telecommunication sector has not only undergone liberalisation reforms across all EU member countries, it also provides sufficient variance in the degree to which reforms have resulted in greater competition and an increase in the number of service providers (European Commission, 2010; Conway and Nicoletti, 2006). In this sense, the European telecommunication sector has experienced a strong wave of liberalisation efforts in the 1990s, aiming at the withdraw of market entry barriers and establishing a common European telecommunication market (Conway and Nicoletti, 2006; see also Clifton, Comin and Díaz-Fuentes, 2007). This was indeed one of the most ambitious liberalisation projects of the European Commission (Belloc, Nicita and Parcu, 2013), which served as a European-wide 'laboratory' for provider choice. However, while Conway and Nicoletti (2006) and their OECD-wide compilation of regulatory indicators indicate massive liberalisation efforts, there remain some variation across countries. This is especially true with regard to effective competition between service providers within national markets. Thus one can see that we have a European telecommunication market that is on the one side highly liberalised, but on the other side differs with regard how those reforms have resulted in greater competition among providers, and choices available to citizens. This makes it an ideal case for our subsequent empirical analysis.

5. Data and Methodology

We use data from the European Commission's Eurobarometer project. Eurobarometer surveys are known for their high quality and methodological rigour in both survey design and data collection. Adopting a multistage, random probability sampling procedure, information is collected, through face-to-face interviews at respondents' homes (GESIS, 2010) – yielding a total of 24,815 respondents. In our study, we use data from Eurobarometer 65.3 on services of general interest (European Commission, 2006). The survey was fielded 2006 in 25 EU member countries. We filtered out those respondents who were not service users and deleted cases with item non-responses. This resulted in a sample of 15,143 mobile service users and 13,422 fixed telephony users.

Dependent variables

In our study, we examine individual level switching behaviour in 2006 within the mobile and fixed telephony sectors of the 25 member countries then part of the EU. Here, Eurobarometer 65.3 contains relevant information on citizens' switching behaviours in both sectors. More precisely, respondents were asked 'Have you tried to or thought about switching your [insert service] provider in the last two years?'. Possible answers were: 1 'Yes, you switched and it was easy', 2 'Yes, you switched but it was difficult', 3 'Yes, you tried to switch but you gave up switching due to obstacles you faced', 4 'No, you did not try because you are not interested in switching' and 5 'No, you did not try because you thought it might be too difficult'. Our interest is in whether public service users have actually switched providers, and therefore we coded this as a dichotomous variable. Respondents that indicated that they had switched providers in the past two years were coded as '1' while the non-switchers were coded as '0'. Overall, 18% of fixed telephony and 25% of mobile telephony users had switched their providers within the period surveyed. Figure 4.1 provides a disaggregated overview of switchers for both sectors and one can clearly see that there are significant differences in switching behaviour across countries.

Potential vulnerability

We argue that citizens' switching behaviour differs in accordance to their vulnerability, which has been argued of being a latent concept – meaning that it is not directly observable. While potential-vulnerability can be operationalised in various ways (see for example OECD, 2008), a low educational attainment has been one of the most remarkable and repeatedly used operationalisations (for example Burden, 1998; Clifton et al., 2011a, 2011b; George et al., 2011; Jilke and Van de Walle, 2013). This is not without a reason, educational attainment represents a particular element of the concept of vulnerability, which is cognitive ability. It largely affects

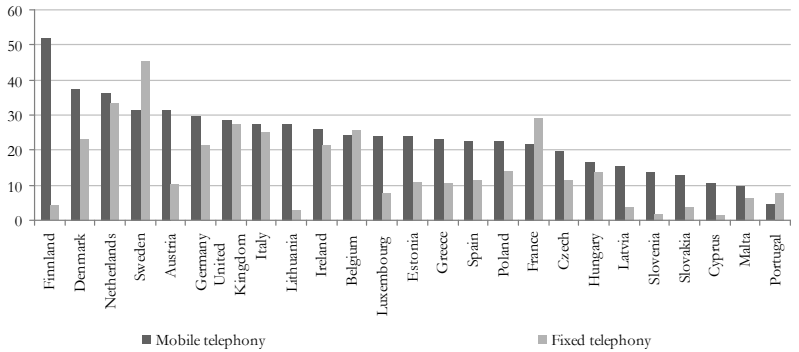

Figure 4.1: Service users' switching behaviour (percentages).
Source: Own calculations using EB65.3 data.

consumers' resources for participating in the market (Hogg et al., 2007). One the one side education impacts the development of skills for consumer empowerment, including the acquisition of information and the knowledge of how to interpret them (Brennan and Coppack, 2008). On the other side, as suggested by Clifton et al. (2011), a low level of formal education is strongly associated with people's more limited resources for processing and evaluating information (see also Dohmen et al., 2010; Hjorth and Fosgerau, 2011). For our study, we look at the cognitive ability element of the concept of vulnerability, as we argue that especially people's cognitive resources affect their choice behaviour. Doing so, we grouped respondents, based on their age when they left fulltime education, into three categories: basic education, secondary education and higher education. Respondents who were still studying were assigned to one of the three categories based on their current age.

Number of service providers

The degree of choice that is available to public service users is measured through the number of service providers within national telecommunication markets. The European Commission provides estimates of the number of service operators within both telephony sectors. However, these numbers are based on different national definitions of which providers to include, and thus do not allow cross-national comparison. Therefore, we established our own values using a common definition of service provider: a public or commercial organisation that provides voice telephony services on a national basis, thereby excluding, for example, those that offer only international calls. Service providers were identified from national network agencies and provider websites through an extensive web-search. This data has been collected by the author. We individually measure the number of providers of mobile and of fixed telephony who had entered the respective telephony market prior to January 2005.

Control variables

We controlled for a number of other socioeconomic variables, namely gender, age, employment status, place of residence and homeownership. Males have been shown to be more likely to switch their service provider and this is perceived to be because of their greater interest in technological innovation (Ranaganathan et al., 2006). Furthermore, we take into account that the elderly are thought be less active in their switching behaviour. Thus we control for respondents' age. Income has been shown to be negatively correlated with the likelihood of switching, as poorer people are more in need of better service offers (Eppling, 2002). Given data limitations, we are not able to directly measure respondents' incomes or wealth status, and instead use homeownership and employment status as proxy indicators. The place of residence should also be critical in providing services of general interest, as it is often argued that rural areas tend to be under-provided (Clifton et al., 2011b). We therefore also controlled for a respondent's place of residence.

We also control for individual perceptions of service delivery, reflected in aspects such as the daily importance of the service, and switching barriers. Earlier studies into telecommunication switching behaviour have indicated the importance of service usage (Ranaganathan et al., 2006), with frequent users being more likely to change their provider. Thus, we control for the perceived importance of the services used daily, using a four-point Likert scale ranging from not at all important to very important. Another important aspect when it comes to switching behaviour in telecommunications is the barriers to switching (Kim et al., 2004; Lee et al., 2001). Hence, we take into account public service user's perceptions of switching barriers, namely their evaluations of the ease of comparing offers from different providers and the extent to which consumer interests were protected. The ease of comparing offers was measured in the original survey using a four-point Likert scale ranging from 'very difficult' to 'very easy'. Additionally, respondents had been asked to assess how well consumer interests were protected within each service market using a four-point Likert scale ranging from 'very badly' to 'very well'.

On the country level, we control for institutional switching barriers, for the average price for making a call, the total number of subscribers, and the market concentration. In terms of barriers to switching, we include measures for both the services being considered. Within national markets, the number portability rate – that is the average number of days it takes to transfer a phone number from one provider to another – is commonly applied as a measure of switching barriers (see European Commission, 2010). Here, we chose to use the official figures for average number portability between two providers, as reported by the European Commission (2008). We also control for the actual price levels of the services. For fixed telephony,

we use the costs of a ten-minute local call. In terms of mobile telephony, we use the average price per minute of a voice communication (European Commission, 2009a, 2010). However, since absolute price levels differ among countries, we have adjusted these prices by weighting them with their respective Purchasing Power Parities for 2006. Further, we recognise that the number of service providers may not only reflect the degree of market liberalisation but also the size of the market. To control for this, we include the total number of network subscribers in 2006 as one of our country-level predictors. To also control for different degrees of market concentration within national markets, we added the Hirschman-Herfindahl index (fixed telephony)/ Concentration ratio (mobile telephony) for service operators to our models.

Modeling strategy

Given the hierarchical structure of our data, individuals nested within countries, multilevel modelling techniques are required since these are able to correct for potential clustering effects and unobserved heterogeneity across countries (Hox, 2002). Moreover, in order to be able to model individual level predictors of a binary dependent variable (in our case, citizens' choice behaviour) and country level individual variables simultaneously, we estimate a logistic random intercept model. In our analysis, we grand mean centre all our continuous predictors such that the intercept can be interpreted as the value (in terms of the used indicators) attached to the average respondent. Our main individual level predictor education has been group mean centred as we are interested in the individual within-country effects of education, and not in structural differences across countries (Enders and Tofighi, 2007).

6. Results

For both mobile (Table 4.1) and fixed (Table 4.2) telecommunication services, we estimate three separate models. In the two tables, we report odds ratios and standard errors (in parentheses) but, because of space considerations, only the results from our main variables of interest are reported here, with complete results being included in Appendix A. The null model reflects an intercept-only model, which helps to assess how much of the variance can be attributed to differences between countries. In a second step, we added all our independent variables to the models, ignoring any potential interaction between the number of service providers and respondents' level of education. In the third model, we added cross-level interaction terms between choice and education. As regards mobile telephony, each model significantly improved its fit over the previous model (Table 4.1). This is

Table 4.1: Modelling citizens' switching behaviour towards mobile telephony services (EU25).

	Model 0	Model 1	Model 2
Intercept	.292** (.036)	.195** (.045)	.195** (.045)
Number of service providers		1.112* (.052)	1.113* (.052)
Basic education (Ref: higher education)		.864* (.059)	.896 (.063)
Secondary education (Ref: higher education)		.917 (.042)	.933 (.044)
Number of service providers X Basic education			.950* (.023)
Number of service providers X Secondary education			.965* (.015)
Variance: country intercept (SE)	.609 (.091)	.435 (.068)	.435 (.068)
Deviance	16,359.92	16,016.05	16,009.39
Interclass correlation	.101	.054	.054
N (Individuals)	15,143		
N (countries)	25		

*Note: Results of control variables are provided in the Appendix A; Odd ratios with standard errors in parenthesis are reported; significance levels: *p<0.05; **p<0.01.*

Table 4.2: Modelling citizens' switching behaviour towards fixed telephony services (EU25).

	Model 0	Model 1	Model 2
Intercept	.131 (027)	.048** (.016)	.084** (.043)
Number of service providers (log)		2.888** (.413)	1.716 (.678)
Basic education (Ref: higher education)		.693** (.054)	.553* (.128)
Secondary education (Ref: higher education)		.845** (.048)	.824** (.054)
Number of service providers (log) X Basic education			1.161 (.168)
Number of service providers (log) X Secondary education			1.015 (.021)
Variance: country intercept (SE)	.995 (.149)	.283 (.056)	.274 (.053)
Deviance	11,355.11	11,119.03	11,115.81
Interclass correlation	.231	.024	.022
N (Individuals)	13,422		
N (countries)	25		

*Note: Results of control variables are provided in the Appendix A; Odd ratios with standard errors in parenthesis are reported; significance levels: *p<0.05; **p<0.01.*

reflected in the significant decrease in deviance (−2 Log likelihood) when applying a likelihood-ratio test. In the mobile telephony analysis, the intercept-only model revealed an interclass correlation of .101 indicating that roughly 10% of the total variance can be attributed to country differences. Our final model explains 54% of the variance that lies between countries.

Many of the control variables made a statistically significant contribution to our models, and confirmed the expected effect directions. For example, females are also less likely to switch, so as those respondents that own a house, or are older than 69 years. Respondents that place a great daily importance on their mobile service are more likely to switch, so as those public service users that think there is no good protection of their consumer interests. Moreover, our main predictors of interest have the expected effects: being comparatively low educated decreases the likelihood of switching mobile service providers. Considering our level-2 predictors, our findings show that when the number of mobile service providers increases, the probability of switching also increases. This supports the notion that a greater choice does lead to a situation where service users are more likely to opt for another provider. We also find that a greater level of market concentration is associated with lower switching rates. The other country-level control variables were not statistically significant, but the indicated effects were in the expected directions.

Turning to the hypothesised interaction between education and the number of service providers, we find that our interaction term between being low educated and the number of service providers turns statistically significant. Thus, there is initial evidence for an interaction between education and greater choice. We further examined this relationship and calculated the marginal effects of basic education on switching (compared to a high level of education), contingent on the number of service providers, keeping all the other predictors constant at their mean values (see Brambor et al., 2006). The resulting graph (Figure 4.2) reveals an interesting picture: the initially positive marginal effect on switching turns negative with more than five providers, but if we instead consider the 95% confidence interval then the band includes zero up to eight providers. This means that typically there is unlikely to be a negative effect of being low educated on the probability of switching within countries where there are less than eight mobile providers. However, if there are more than eight providers, the marginal effect of a low level of education on switching is clearly negative. Moreover, the revealed effect size is nontrivial; we find in national markets with 13 providers that individuals who are low educated are 9% less likely to switch when compared to their better-off counterparts. The 95% confidence bounds range between 3% and 15%. On the other side, in markets with 8 providers or less, there are no significant differences between low and well edu-

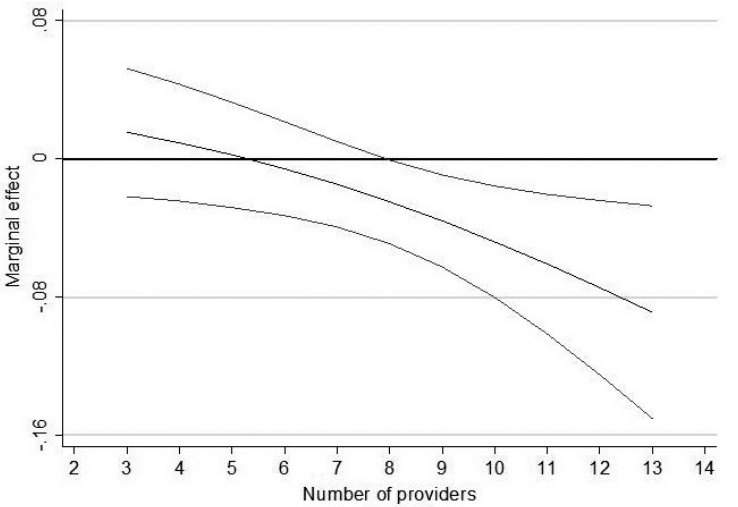

Figure 4.2: Marginal effects of being low educated on the probability of switching mobile services contingent on the number of providers (95% confidence intervals).

cated respondents. This shows that once the number of service providers reaches a threshold of eight service providers, less-educated service users become less likely to switch.

We now turn to the results of our estimations for fixed telephony services (see Table 4.2). Here, the intercept-only model has an interclass correlation of .23, which means that 23% of the total variance in switching behaviour is on level-2. Overall, model 1 significantly improves its fit over the intercept-only model by including additional parameters. However, our interaction model (Model 2) did not significantly improve its fit over model 1 – the difference between the deviances of the two models is too low to satisfy conventional significance levels. This indicates that our interaction terms fail to make a valuable contribution to explaining users' switching behaviour. In terms of our control variables, some were found to be statistically significant, with effects in the anticipated directions. For example, respondents that regard the consumer interest protection in their country as bad are less likely to switch providers. Also, the elderly and those who are living in rural areas are less likely to exercise choice. Turning to our main predictors of interest, as expected, being less-well educated decreases the likelihood of switching service providers. This is in line with findings from the mobile telephony sector.

On the country level, including the number of service providers did yield statistically significant estimates. Although the number of service providers has a positive effect on the likelihood of switching, it has a log-linear form, suggesting diminishing marginal effects of the number of alternatives on citizens' switching. Overall, this finding is in line with our results from the mobile telephony sector. Looking at country-level control variables, only our predictor of market concentration was found to be statistically significant. As the results for the mobile telephony sector do, this suggest that greater market concentration is associated with lower switching rates. The other controls do not reach conventional levels of statistical significance.

Evaluating our research question, we checked for a potential interaction effect between basic education and the number of service providers. While our statistical tests showed that education had indeed an individual effect on the likelihood of exercising choice, the data failed to identify any joint effect. Taking into account the identified interaction between education and the number of providers within the mobile telephony sector, our results regards the equality of liberalisation reforms are mixed. An interesting question is why, with less well educated service users, an increasing number of service providers has a negative effect on the education-switching relationship only within the mobile sector? A first observation is that the mobile telephony market is much more strongly driven by new technological innovations that require greater capabilities to follow than the fixed telephony market, which remains a relatively simple service. As such, the mobile service market can be regarded as a complex environment in which to make switching decisions, whereas, as observed by Iyengar and Lepper (2000), information overload is less likely to take place within rather simple environments. Moreover, the mobile telephony market is much more prone to competition with considerably more service providers on average (mean 7.1, standard deviation 2.9) than the fixed telephony market (mean 3.4, standard deviation 2.4). The negative effect of too great a choice therefore may only unfold if the respective service sector is characterised by a strong market orientation with a comparatively large number of options. In the mobile telephony sector, this threshold seems to be at eight providers.

7. Discussion and Conclusion

The introduction of choice and competition into public service delivery rests on the assumption that overcoming state monopoly-led provision of public services would result in a more efficient process of service delivery, and an increase in citizens' welfare (Le Grand, 2007; Ostrom and Ostrom, 1971; Savas, 1987). For instance, this is done through shifting the autonomy for decision-making from the state

to the citizen by creating markets for public services and letting service providers compete for customers. Public service users send market signals to suppliers by complaining, or switching service providers. As a result a better match between citizens' demands and preferences, and the price and quality of the offered services would emerge. However, in our analysis we have shown that potentially vulnerable and non-vulnerable groups of citizens do not send market signals in the very same manner to providers under different levels of choice. In terms of equality, we have found that increasing the number of choices that are made available to citizens appear to work better in some public services as fixed telephony than in others as mobile telephony, mostly due to characteristics as the competitiveness of the market, and the overall service complexity. However, the question is not whether to open public service delivery for competition and provider choice, or not, but rather how much choice works for a given service. Once 'too much' choice is made available a choice-gap is likely to emerge.

One has to note that our results indicate that liberalising public services does not per se negatively influence the switching behaviour of potential vulnerable groups but that, for this to occur, a certain threshold of provider choice must be exceeded. The circumstances under which the introduction of choice negatively impacts on the switching decisions of the potentially vulnerable are not clear cut and may vary across different public service sectors. We have identified criteria that, if satisfied, could result in liberalisation reforms creating a 'choice-gap'. This can occur if the public service sector exhibits a strongly liberalised and competitive environment with a high number of providers. Further, we suspect that the less complex an actual service is, the higher the number of providers needs to be before the negative potential becomes a reality. However, these criteria should be subjected to further testing by future research.

There are, of course, some limitations of our analysis which we believe could be addressed by future research. In terms of the generalisability of the results, we cannot confidently claim that similar effects would be found for other types of public services such as employment services, gas, or health care. Thus, future studies may look whether our results hold true for other types of public services. Furthermore, the availability of choice, or number of service providers within national markets, is a key goal of the EU liberalisation agenda and has been argued to be positively related to the degree of competition within countries. However, the market share of single providers may differ across countries and thus not necessarily equate with the degree of competition. Our data does not allow to examine the overall market share of single providers, as such data is not publicly available (only on the level of operators). Instead, this study focuses on the availability of choice to public

service users, and not competition per se, while controlling for the actual market concentration (using the Herfindahl-Hirschman index/ concentration-ratio for service operators). We argue that it is important to estimate the relationship between the number of available alternatives in national markets and vulnerability, because a great variety of service offers was a key policy goal in the implementation of liberalisation policies across the EU. Thus how service users respond to an increasing number of telephone providers within national markets is an important question of theoretical and practical pedigree. Furthermore, the exclusion of a measure for public service users' income may bias our results. To account for this, we used income-related (state-of-the-art) proxies as control variables, including respondents' homeownership and employment status. We also need to acknowledge the cross-sectional nature of our data. While we account for wide range of control variables at both, the individual and country level, we cannot confidently rule-out that our findings may be affected by other unobserved factors, or reverse-causality. Instead, what this study can do is to pinpoint an association between vulnerability and switching behaviour, which increases with a growing amount of options to choose from. Future studies are well advised to collect time series data for citizens' switching behaviour to cross-validate our result. For now, our study results clearly show the interesting relationship between citizen vulnerability and decision making in public service markets.

Our study has some important implications for policy makers and regulators. We have shown that an increased number of options to choose from can have heterogeneous effects on the switching behaviour of vulnerable and non-vulnerable service users. However, for this effect to unfold, a certain threshold must be reached (in our case eight providers), and the market needs to be strongly de-regulated. This may suggest to limit the number of licences awarded to providers in de-regulated markets. However, instead we would argue that potential vulnerable services users should rather receive greater attention in consumer protection policies. Most policy attention has gone into reducing switching costs (e.g. number portability rates), instead search costs seem to be as important. Thus establishing independent agencies that provide easily available information on service offers could be one possible responses to an increased market complexity, and help to reduce the 'choice-gap'. Another possible way to account for an increasing inequality in citizen choice behaviour could be service obligation contracts that protect the vulnerable. Such contracts outline arrangements between non-public service providers and the regulating public bodies (for an overview see Cheung, 2005). They could, for example, include guidelines to strengthen the market situation of vulnerable customers across the EU. Additionally, organising collective switching schemes, as in the electricity market in the UK (see for example UK Department of Energy and

Climate Change, 2013), could also be one interesting way to overcome inequalities in citizens' choice behaviour. In such switching schemes municipalities select providers for a large group of service users on a tender basis and then collectively switch to the one with the best service offer. Enrolment in such schemes is relatively easy and straightforward. However, whether this can accurately reflect heterogeneous consumer preferences – especially in the mobile telephony market – is not clear and should be the subject of future studies.

CHAPTER 5

Two Track Public Services? Citizens' Voice Behaviour towards Liberalised Services in the EU15

This chapter has also appeared as Jilke and Van de Walle (2013)

1. Liberalised Services and Citizens' Voice

Turning public service providers into corporations that employ private sector management techniques has been on the agenda of policy makers for some time (Clifton, Comín and Díaz-Fuentes, 2007). Following large-scale reforms, inspired by European integration and liberalisation, the publicness of many so-called services of general interest[23], such as public utilities, public transport, postal services etc. changed (Héritier, 2001; Van de Walle, 2008). The purpose was to turn state monopolies into competitive private or semi-private providers. For citizens, this meant they became to be seen as vocal, empowered consumers of public services, rather than mere legal subjects. According to this model, citizens should be viewed as customers who make well-informed choices, with all rights and entitlements to consideration and service that this entailed (Aberbach and Christensen, 2005; Clarke and Newman, 2007).

Recent government initiatives in this regard have been increasingly focused on providing greater opportunities for citizen voice[24]. The introduction of more extensive voice mechanisms (along with the possibility of provider choice) aims at improving the responsiveness and subsequently the performance of public services by emphasising the central role of the citizen in the process of service delivery (Le Grand, 2007). They are also important constituents within the liberalisation of services of general interest debate in European public sectors (Clifton and Díaz-Fuentes, 2010; Clifton et al., 2005; Vael et al., 2008).

Many authors have discussed the anti-democratic implications of these kind of reforms (Behn, 1998; Box et al., 2001; Gottfried, 2001), especially their propensity to establish a 'supermarket state' model, where the wealthiest, best-informed and most assertive customer get the best quality service (Olsen, 1988; Christensen and Lægreid, 2002). In line with this, services of general interest reforms have been severely criticised, as it is thought this shift has disadvantaged potentially vulnerable citizens (Clifton et al., 2011) and created a 'two-track' public service. This research note provides a very first empirical outlook on this assumption by analysing Eurobarometer public opinion data on stated voice behaviour (complaints) towards

[23] For the sake of clarity we do summarise service of general economic interest as well as service of general interest under this term. Their legalistic distinction has been extensively discussed elsewhere (Prosser, 2005).

[24] In this research note we focus on the formalised individual element of voice within the service delivery cycle, what Dowding et al. (2000) call individual voice, namely complaining about any aspect of used services.

liberalised public services in 15 member countries of the European Union (EU) for the years 2000 and 2004[25].

The shift from collective-oriented models to individualised customer-based models have been studied from a variety of academic disciplines, and often highlight the managerial imperatives at the heart of these reforms (Clarke and Newman, 2007; Learmonth and Harding, 2006). More recently, we have also seen increasing attention to studying public attitudes towards public services (Fiori et al., 2007; James, 2009; Van Ryzin, 2006) and towards the effects of liberalisation in certain sectors (Brau et al., 2010; Bacchiocchi, Florio and Gambaro, 2011; Clifton et al., 2011; Fiorio and Florio, 2011). Examples include studies looking at cohort changes in satisfaction with health care (Adang and Borm, 2007) or studies looking at public preferences for or against public or private provision of public services (Wendt et al., 2010). In this research note, we want to build on these initial studies.

The structure of this research note is as follows: first we introduce our research question, so as data and methodology. Findings from our empirical analysis are presented consecutively where we assess whether an association can be found between citizens' voice behaviour and their socio-economic status, and if we can establish a link towards the liberalisation of services of general interest. Following, we discuss our results, limitations of the study and draw some conclusions for future research.

2. Research Question, Data and Methodology

We aim at assessing the relationship between citizens' socio-economic status and their stated complaint behaviour towards liberalised services of general interest in 15 EU member countries for the period between 2000 and 2004. Hence we test whether there are differences for different socio-economic groups in exercising voice towards any aspect of a received service. We are particular interested whether potential differences between contrasting layers of society increased (or decreased) over time, and if so, whether this process can be attributed to the liberalisation of services of general interest. Furthermore, we assess our findings in the light of data and item availability and conclude with indications for the future research agenda in this field of study.

[25] Within this time span, liberalisation reforms in the EU15 were still ongoing (Prosser, 2005) and the overall regulatory status (measured through the OECD's ETCR indicators) between 2000 and 2004 for all those countries differs significantly (Conway and Nicoletti, 2006).

We estimate a binary logistic regression model using Eurobarometer data between 2000 and 2004. In general, Eurobarometer public opinion surveys are commissioned by the European Commission twice a year since 1973. They contain approximately 1,000 respondents in each country[26]. National samples are drawn using a multi-stage random probability sampling design from the total population aged 15 and above. The data is weighted in proportion to its share in the total population of the EU15 member countries, aged 15 and over. These adjustments are based on EUROSTAT population figures and include post-stratification sample weighting factors (GESIS, 2008).

Between 1997 and 2008, six Eurobarometer rounds have been devoted to citizens' perceptions on (public) services (47.1 in 1997; 53.0 in 2000; 58.0 in 2002; 62.1 in 2004; 63.1 in 2005; 63.1 in 2006), plus two flash surveys[27] – one on service quality (flash 150) in 2003 and one on switching service providers (flash 243) in 2008. For our analysis, we merged the most relevant datasets (2000, 53.0 and 2004, 62.1). In total, they provide a sample of 31,429 respondents for all EU15 member countries – due to missing values, our sample dropped by 3% to 30,488 respondents.

Dependent variable

Within the analysis, we will focus on the impact of citizens' socio economic factors towards their stated voice behaviour, using the following item:

"In the last twelve months, have you personally made a complaint, either to any complaint-handling body (ombudsman, regulator, consumer association, industry body, etc…) or the service provider about any aspect of […]?".

As our dependent variable, we constructed a binary variable (has complained about one of the services/ has not complained at all; see table 5.1), incorporating the following liberalised services: electricity, gas, fixed phone, mobile phone, water, postal services, rail and local transport, by simply adding them up. We use this composite measure of voice in order to reduce effects of missing cases within certain service sectors[28], as well as to increase the total number of complaints within our analysis. The advantage of this procedure is that it makes the analysis less sensitive and more robust. The other side of this rather pragmatic approach is that we aggregate

[26] With the exception of Germany (2,000 respondents), Luxembourg (600 respondents) and the UK (1400 respondents).

[27] While standard and special EB rounds are conducted through face-to-face interviews, flash surveys collect information by telephone interviews.

[28] The amount of missing data and non-users within single service sectors compromises almost 10% and above.

Table 5.1: Descriptive Statistics.

	Mean	SD	Min, Max	N
Dependent variable				
Voice	.17	.377	0, 1	30,570
Independent variables				
Education	1.98	.757	1, 3	31,120
Age	2.75	1.072	1, 4	31,429
Control variables				
Country	–	–	1, 15	31,429
Year	.49	.500	0, 1	31,429
Service quality	11.42	4.316	1, 32	30,479
Gender	.48	.500	0, 1	31,429

Source: Eurobarometer 53.0, 2000; Eurobarometer 62.1, 2004.

experiences within specific service sectors. However, these sectors do have indeed different characteristics within and across single countries – we will further discuss this point in the discussion of results.

The 2000 (53.0) dataset is the first Eurobarometer round that contains information whether respondents have actually submitted a complaint. It has been incorporated in the same way repeatedly until 2004 (62.1). In 2006 (63.1) the question was asked in a more fragmented manner, as it was changed from a binary answer to four answer possibilities, probing for different kinds of institutions complaints could be submitted to. This affects comparability negatively, as it reduces the overall probability in selecting "no" as an answer by decreasing the share of the "no" category from 50% to 25%. Furthermore, in 2006 the time span was extended to 24 months. Another difference which was observed between the Eurobarometer rounds in 2000 and 2004, was that in 2004 the question on voice was asked exclusively to service users (those respondents that had stated that they actually use the service and have access to it), while in 2000 all respondents were included. In order to reduce the effect this procedure might have on the results obtained, we filtered those respondents in 2000 who had stated that they have no access to the respective service – figures on service use were not available at that time. This still may have a minor impact on the results obtained.

The cross-national reliability of the survey item has been tested by comparing item non-response size and characteristics of sub-components of our dependent variable (specific service sectors) per country. Results reveal no significant differences across countries for those respondents opting for the "don't know" answer.

Independent variables

For our independent variables we utilize available socio-economic indicators as proxies for potential vulnerability (see also Clifton et al., 2011), namely the respondents' age and educational status – income was not included in the initial Eurobarometer database. We constructed four age groups: 15-24 years; 25-39 years; 40-54 years; and 55 years and older. As regards educational status, we grouped respondents in accordance to their age when they left fulltime education into three categories: basic education (<15 years), secondary education (16-19 years) and higher education (>20 years). In order to minimise effects, respondents which were still studying were assigned to one of the three categories in correspondence to their age.

Additionally, we included two interaction terms (Aiken and West, 1996) for Year (2000=0, 2004=1) * Education and Year * Age. By this, we aim at assessing whether there is an interaction between time and potential vulnerability when it comes to exercising voice. In other words, the first term measures whether educational attainment over time increases or decreases the probability of submitting a complaint. The second term assess whether the same holds true for age. Following this procedure, we want to find out how a potential equality gap in terms of exercising voice between different socio economic groups develops over time.

We seek to control for important factors that have an influence on citizens' voice behaviour. First we control for country effects by constructing dummy variables with Ireland as reference, which has, in terms of voice, the closest mean to the grand mean. Furthermore, we controlled for perceived service quality of received services, where we constructed a composite measure by adding perceived service quality variables[29] from all services included. On top we included the respondents' gender and the survey's year as additional control variables. The Eurobarometer dataset does not provide other potentially important control variables such as expectations towards service quality – we will take this issue into account in the discussion of the results.

3. Patterns of Citizens' Voice Behaviour in the EU15

We first examined developments over time for the share of respondents that have actually submitted a complaint for all available service sectors in the EU15 (Figure

[29] The item was phrased as follows: "Overall, what do you think of the quality of […] that you use? Would you say it is very good, fairly good, fairly bad, or very bad?". It has been incorporated for all service sectors used within this study. We have then simply summed-up all answers to a single variable, which we have labelled "service quality" (see also Table 5.1).

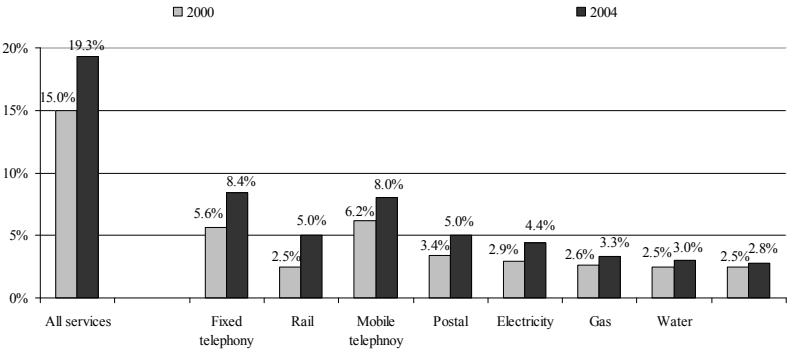

Figure 5.1: Complaints made in the last 12 months (EU15).
Source: Own calculations based on Eurobarometer 53.0, 2000; Eurobarometer 62.1, 2004.

5.1). Here an increase in complaints – indeed, in different degrees – was apparent in all service sectors. However, we can clearly identify differences across sectors. While there has been a rather strong increase in complaints in the telephony and rail sectors, the remaining sectors experienced only some minor changes of less than two percentage points between the years.

In this regard, we estimate a binary logistic regression model (Hosmer and Lemeshow, 2000) with voice behaviour as dependent variable. In general, we observe a relatively low Pseudo R^2 and Nagelkerke R^2, showing that our model does not fully explain citizens' voice behaviour, however, that was not our intention. Moreover, we do find a good model fit for our model.

Findings reveal that country effects are apparent which suggests that there are considerable differences in exercising voice between countries (see table 5.2). Most controls are statistically significant and pointing in the expected direction. Service quality of the received service, so as developments over time, does play a statistically significant role in this regard. The only control variable that had no effect on voice behaviour was gender. Regarding our independent variables, both predictors, education and age, are statistically significant. It shows that less educated respondents are less likely to submit a complaint when compared to those that exit the educational system at a later stage. The magnitude of the odds in this regard is stronger for low educated than for medium educated respondents. As regards age, we find middle aged respondents to be more likely to exercise voice when compared to those 55 years and older. Here again, we can find some differences within the odds, showing that respondents between 25-39 years are the 'strongest' group, followed by those 40-54 years old.

Table 5.2: Binary logistic regression for reported voice behaviour.

	EU 15	
	B	Exp(B)
Control variables		
Fixed effects (omitted)		
Year 2004 (Ref. 2000)	.287	1.332**
Service Quality	.079	1.082***
Gender (Ref. male)	–.049	.953
Independent variables		
Education (Ref. high)		
Education low	–.636	.529***
Education medium	–.279	.757***
Age (Ref. 55+ years)		
15-24 years	.416	1.516***
25-39 years	.393	1.481***
40-54 years	.412	1.509***
Interaction Terms		
Year X Education low	.434	1.544***
Year X Education medium	.024	1.024
Year X 15-24 years	.106	1.112
Year X 25-39 years	.229	1.257*
Year X 40-54 years	–.081	.022
Constant	–2.626	.072***
Nagelkerke R2	.084	
Pseudo R2	.050	
Correctly predicted	82.9%	
N	30,488	

*** p<,001; **p<,010; *p<,050

The inclusion of interaction terms did improve the explanatory power of the statistical model and had no significant effect on our independent variables. Here we can see that education had a positive but small effect on our dependent variable over time, which may also imply that low educated people have complained more frequently in 2004 than in 2000. This is contrary to our expectation that education decreases the odds of complaining over time. As for age, we find that those respondents 25-39 years old are complaining more often over the years when compared to those 55 years and older. In other words, this point in the direction that the effect of education and age on exercising voice, namely that less educated and elderly are less likely to do so is differing. While being low educated seems not to decrease the

probability of exercising voice over time, being 55 years and older did. However, the observed effect for age is rather small.

Our findings only partly support our expectation that socio-economic factors did have a negative impact over time on voice behaviour towards liberalised services. While education has not such an effect on citizen complaint behaviour, age has. Hence our data suggest that the gap in submitting complaints between the young and vital, and the elderly increases over time. As regards educational status, that opposite seems to hold true, meaning that the gap between different educational groups decreases over time. However, these results should be regarded as provisional for various reasons, on which will elaborate in the following.

4. Discussion and Future Research

So far our findings do indicate an equality gap between different layers of society when exercising their voice by submitting a complaint. More precisely, results suggest that lower educated citizens are less likely to submit a complaint when compared to those with a higher education. Age has a considerably weaker effect on the likelihood of complaining. These findings are in line with previous research (Dowding and John, 2008; Thomas and Melkers, 1999; Thomas, 1982). Furthermore, we provide first support for the assumption that the gap between the young and the elderly does increase over time. Education, against our expectations, did not have a similar effect. In contrary, being low educated slightly increases the likelihood of submitting a complaint between 2000 and 2004, which suggest that the gap between different educational groups was decreasing in that particular time period. In this regard, we have to keep in mind that these are only overall tendencies that do not necessarily reflect sectoral developments which might vary accordingly.

Our analysis may suffer from a number of potentially biases that need to be discussed. First we had to aggregate voice behaviour for specific service sectors to a single item in order to produce reliable results. On the one side this procedure solves a methodological issue, but on the other it raised a new problem. Reforms within different countries and different service sectors are diverse in terms of reform depth and speed (Conway and Nicoletti, 2006). Therefore, our composite measure of voice is not able to trace developments within specific sectors, but it may provide a first overall direction (see also Ferrari, Pagani and Fiorio, 2010 for a comparable approach using a composite measure across four different sectors for satisfaction with services of general interest) for more in depth research in the future. Second, attributing our findings to the process of service liberalisation proves to be difficult

for two reasons: we did provide no real ex-ante/ ex-post comparison of reforms, and our timeline of four years is rather short; different degrees of liberalisation within different sectors cannot be adequately attributed to our aggregated measure of citizens' voice.

In this light we cannot be fully confident that our findings stay unchanged when adding additional controls (such as expectations), or assessing developments in single service sectors. Hence, our findings provide only a tendency in this regard. This exemplifies the fact that citizen complaint behaviour in the light of service liberalisation proves to be a complex issue that needs to be assessed at a more disaggregated level. One finding, however, that seems, more robust is that low educated and elderly people are indeed significantly less likely to compliant about their services. In this regard, we regard our results as a first stepping stone in contributing to a wider puzzle: the effects of service liberalisation reforms on citizens' attitudes and behaviours in European public sectors.

Further research efforts in this strand of study need to take various factors into account. As we have shown, citizen complaint behaviour might be affected by socio economic factors. In this regard, age and education are certainly of importance, but there are maybe other socio-economic variables that also influence ones' complaint behaviour, such as wealth, or social-class, or even more general (and perhaps underlying) sets of values. In the future, those set of variables need to be taken into the equation when estimating voice behaviour. Furthermore, relying on secondary data we estimated our model for stated and not observed complaints. Future work may also make use of data on observed complaint behaviour in order to capture actual rather than just reported behaviour.

Moreover, we need to expand our focus by looking not only at short time spans, so as individual factors but also at institutional factors such as different regulatory regimes within a longer period of time. Hence we are in need for further research on the impact of service liberalisation on voice behaviour taking long-term developments or ex-post/ ex-ante comparisons of liberalisation reforms within countries and specific service sectors into account, ideally using a longitudinal multilevel modelling strategy. Therefore, future investigations may use additional data sources since Eurobarometer is limited in terms of comparison over time, and changes of the wording within the Eurobarometer questionnaire do not permit comparisons beyond 2004. In the case that such data is not available, a sole multilevel model, focusing on the regulatory status of specific service sectors, would also produce potentially interesting findings.

CHAPTER 6

Vulnerable Citizens in Public Service Markets after Regulatory Reforms: Towards an Affordability Gap?

This chapter is a refined version of Fernández-Gutiérrez, James and Jilke (2013).

1. Introduction

During the last decades, European Union (EU) countries experienced deep changes in the regulation of public infrastructure services through market-oriented reforms. These reforms, especially the liberalisation of public infrastructure markets, were key parts of the EU Single Market (Clifton et al., 2006). A central aim of the reforms and subsequent regulation of markets has been the promotion of the consumer interest (Cseres, 2008). It was argued, that in competitive markets with an increasing number of providers, rational consumers would be able to have enhanced opportunity for choice between service providers and would subsequently seek to maximize their utility through market signalling (i.e. switching between, or complaining to service providers).

However, once markets became increasingly complex after liberalization processes, it is not so clear whether all citizens in their role as consumers, independently of their background and characteristics, were equally able to take optimal decisions. This raises the important question of whether all citizens have benefited from the increasing choice opportunities in these liberalised public service markets (Clifton et al., 2011; Jilke, 2014). Citizens' socio-economic status is an important factor with "stronger" and better positioned enjoying more information and higher capacity for interpreting it (Ayeni, 2000). This possibility arises that disadvantaged and vulnerable citizen-consumers are left behind in the market-oriented provision of public infrastructure services. This is a major issue for EU regulatory policy because the services are explicitly considered by the Treaty of Lisbon as key elements of social and territorial cohesion, solidarity and equity.

After detecting citizen-consumers' problems related to market malfunctioning, European regulatory policies have turned to an increasing attention to the consumer perspective. This attention is key for increasing support for the Single Market and for addressing concerns about the economic and social effects of liberalisation (Howarth, 2008). Inclusive benefits for all segments of society is important to the functioning of a European single market and the legitimacy of its institutions. However, the analysis of the functioning of public infrastructure services after market-oriented reforms has, so far, mostly focused on the supply-side perspective (Clifton and Díaz-Fuentes, 2010). In this study, we look at the demand-side by providing an analysis of the impact of market-oriented provision of services to different segments of citizen-consumers, which has important implications for the theory of public service provision and EU regulatory policies. Doing so, we analyse the relationship between the level of competition in public infrastructure markets and differences in the affordability of these services between potentially

vulnerable citizens-consumers and their better-off counterparts. In other words, we look whether the extent to which there exist disparities between different layers of society in terms of the affordability of infrastructure services, are larger, or smaller, in national markets with different levels of competition and choice between service providers. For this purpose, as explained in detail in section 3, we focus on educational attainment as a socio-economic variable representative of the capabilities and social networks essential for accessing and processing the information for decision-making (George et al., 2011; Hogg et al., 2007), and thus representative of citizen-consumers' potential vulnerability. We estimate multilevel models on individuals' self-perceived affordability of public infrastructure services in all the EU-25 countries. The models allow testing of cross-level interactions between levels market competition whilst also controlling for other relevant socio-economic and country-level characteristics. We focus on two essential public infrastructure services: electricity and fixed telephony that have both experienced market-oriented reforms to a high degree (Levi-Faur, 2004). In addition, they have two important characteristics crucial for the analysis: firstly, both services were widely used before the reforms; and secondly, they show significant differences among EU countries in the level of competition in the markets, allowing for differentiation through this variable.

We find that, while less educated citizen-consumers tend to experience lower levels of affordability of the services under analysis than those better educated. However, this gap is not significantly affected by markets' competitive structure. Instead, where this competitive structure is translated into a higher frequency of switching, the affordability gap between different socio-educational layers tends to be lower and eventually diminishes. We suggest that infrastructure providers more closely match their service offers with citizen-consumers demands in markets with high national switching rates. Or, in other words, once a critical mass of citizen-consumers switches among service providers thereby creating market pressures from the demand-side, this creates positive externalities for all citizen-consumers, including those who are potentially vulnerable in the marketplace.

The next section of this study describes the theory and policy issues of market-oriented provision of public infrastructure services from the perspective of citizens as consumers, and discusses the issue of vulnerable consumers. This section sets out the hypotheses drawn from the theory. The third section describes the data and empirical approach to address the hypotheses, combining survey data from citizens' experiences with and attitudes towards public infrastructure services and their socio-economic characteristics, with data on market structure and choice at the country level. The fourth section presents the findings of this analysis. Finally,

the fifth section sets out the conclusions and draws out their implications for regulatory policies.

2. Reforming Public Infrastructure Services for All Citizens?

Major regulatory changes to public infrastructure services in the EU – staring in the 1980s – aimed to infuse market forces into the provision of these services. Telecommunications was the paradigm of these market-oriented reforms (Schmitt, 2013), as well as the most advanced sector in terms of supranationalism in its regulatory regime (Levi-Faur, 2004). However, variegation in the extent of competition introduced, offers an opportunity to design research analysing the relationship between competition, choice and the self-perceived affordability of vulnerable citizen-consumers. In telecoms, there was a multi-speed and differentiated process of liberalisation among EU countries (Humphreys and Simpson, 2008; Thatcher, 1999). Further differentiation existed in the energy sector, where a compromise for an incrementally and moderately market-opening formula was adopted (Héritier, 2001). As a result, some EU countries moved quickly towards energy liberalisation, while effective competition remained rather limited in other member states (Howarth and Sadeh, 2010).

Prior to those reforms, energy and telecommunications were organised as public monopolies or publicly licensed private monopolies, which aimed to guarantee equal access for all citizens (Héritier, 2001; Humphreys and Simpson, 2008). But from the point of view of service users, this regime was criticised for their failure to identify and meet citizens' demands for innovation, modernisation and expansion (Héritier, 2002; Thatcher, 1999). This served to justify market-oriented reforms, which by removing barriers to entry, increasing the number of providers and allowing citizen-consumers to freely choose among them, aimed to lead to cost effectiveness, lower consumer prices and improved quality and variety (Héritier, 2001). As noted by Bartels (2013), the reforms had not only economic motivations, but also aimed to enhance the control, power and choice of service users. This, according to the neoclassical economic view of consumption, expected that all the citizens in their role as consumers were able to benefit in the resulting competitive market environment, by making rational choices among multiple providers and options existing. Thus their consumer welfare was expected to increase (Cseres, 2008).

A key aspect of consumer welfare is the affordability of services. The price of a used services is of universal interest and of particular importance to vulnerable citizen-consumers because they cannot afford to ignore this dimension of the service. The

subjective affordability of services is, by common agreement, an important component of overall welfare and has previously been noted as worthy of study in its own right (Bacchiocchi et al., 2011; Clifton et al., 2011, 2014; Fiorio and Florio 2009, 2011; Florio, 2013). However, from the beginning concerns were apparent about the impact of market-oriented reforms from the point of view of citizen-consumers. Infrastructure services are, in most of the EU countries, considered "public services", understood not in terms of ownership but of the objectives inherent to their supply, addressed to the public interest or the general interest (CEEP, 2010; Van de Walle, 2009). As a result of their key economic and social role, these services are subject to public service obligations, representing general-interest objectives as universality, quality, affordability, accessibility, the protection of consumers and their role in promoting equity and social cohesion. As described by various commentators (e.g. Clifton et al., 2005; Clifton and Díaz-Fuentes, 2010; Héritier, 2001; Van de Walle and Hammerschmid, 2011), critics of those market-oriented reforms highlighted that in a competitive market environment, private providers would pursue commercial interests over general-interest objectives, and would not necessarily provide a uniform set of benefits to all citizen-consumers. Considering that in the Single Market Programme (1987) and the Maastricht Treaty (1992), general-interest objectives were primarily regarded as an obstacle to market integration, these concerns led to serious problems of legitimacy to the EU (Clifton et al., 2005). As a response, so-called "re-regulation" was introduced, subdivided into rules to make markets work, prevent market dominance and boost competition, but also to correct the outcome of market processes so that social and political objectives may be met (Héritier, 2001; see also Wright, 2009 for the UK health sector). Trying to balance both aims, the EU concept of Services of General Economic Interest (SGEI) encompasses the services object of this study: services of an economic nature subject to competition and the market rules, to which citizens assume a role as consumers, but also subject to public service obligations under a general-interest criterion due to its essential character for citizens (European Commission, 2003; Bauby, 2008). With this approach, the EU officially recognises the key role of these services in strengthening solidarity, equity and social cohesion, as one of the cornerstones of the European social model (European Commission, 2004).

Due to the lack of consensus, public service obligations and issues related to social regulation of SGEI have been translated to the national level or should be promoted by the European Commission (EC) at the EU level but using "soft" instruments (Clifton and Díaz-Fuentes, 2010). This has favoured a change in the focus of EU regulation of public infrastructure services from the citizens' rights to the consumers' interest. From this paradigm, the EU has focused on the evaluation of services' performance and accumulating knowledge on consumers' subjective evaluations

of services (Clifton and Díaz-Fuentes, 2010; Clifton et al., 2012; Vigoda-Gadot et al., 2010). According to Howarth (2008), through emphasising the consumer interest, the EU aims to reflect its attention not only to the market integration and liberalisation dimensions of the Single Market, but also to social, environmental and cohesion dimensions, on which public infrastructure services remain key.

However, the evaluation of the sectors considered key for the Single Market has shown serious problems of market malfunctioning in the case of public infrastructure services (Ilzkovitz et al., 2008). Focusing on citizen-consumers' opinions, most of the problems are concentrated in telecommunications and energy, particularly related to the handling of complaints and to difficulties in comparing prices and switching suppliers (FONDACA, 2008). As noted by Cseres (2008), markets' opening up to competition has not resulted in their expected optimal functioning in terms of consumer benefits. Fiorio and Florio (2009) developed an empirical analysis of the effects of market-oriented reforms on citizens' satisfaction with various services on the EU-15 countries, which obtained mixed evidence. In later specific analysis focused on electricity, Fiorio and Florio (2011) found a positive effect on perceptions of affordability, whilst that of privatisation was negative. For fixed telephony, Bacchiocchi et al. (2011) found, in general terms, that regulatory changes were not positively related to people's affordability evaluations. Overall, the existing literature leads to question the mainstream approach to market-oriented reforms of public infrastructure services.

The EC has recently begun to recognise the insufficiencies of regulatory policies from the supply-side, and is looking for new policies based on the consumers' perspective, in a complementary way to the more traditional competition policies. From this base, new consumer policies focus on helping citizen-consumers' appropriate use of market mechanisms, as providing them better information and facilitate them to exercise their rights to switching and complaining (e.g. Howarth, 2008; Xavier, 2008). In this effort, insights from behavioural economics have been very influential. Behavioural economics focuses on the cognitive biases that affect individual behaviour, leading to bounded rationality and limited selfishness (Mullainathan and Thaler, 2000; Xavier, 2008). Thus, it puts into question the view of citizen-consumers' as completely consistent, rational and selfish decision-makers in which the market-oriented reforms of public infrastructure services were based. These insights are useful to understand the functioning of markets, especially in those cases where choice has been recently introduced, as public infrastructure services. Illustratively, Giulietti et al. (2005) found a barrier to switching in the gas service due to searching costs and the perceived cost of the change, whilst Lambrecht and Skiera (2006) for internet and Wilson and Waddams (2010) for

electricity services observed that consumers' decisions do not necessarily imply minimisation of the tariffs paid. According to Gans (2005), benefits of competition may not occur if consumers do not behave in a perfectly rational manner, or do not have perfect information. Thus, as summarised by Cseres (2008), market-oriented reforms aimed at benefitting citizen-consumers would require that markets are transparent, information costs are affordable and consumers have the competence, capacity, opportunity and motivation to take on the responsibilities shifted from the state to individuals. In this context, European policy makers are interested in how further understanding of consumer decision-making could lead to specific regulatory policies to increase their power, involvement and confidence in markets, thus providing a better exploitation of the potential benefits of competition and choice in terms of consumers' welfare and satisfaction (European Commission, 2008a; 2010).

Insights from behavioural economics are crucial for a better understanding of the factors leading to citizen-consumers' decision-making in the markets. An important insight from behavioural economics is that not only the product matters for consumers' decision-making, but also contexts and individual capabilities (Cseres, 2008). Combining insights on bounded rationality and limited selfishness with analysis on the influence of the social context on decision making, Clifton et al. (2011; 2014) and Jilke (2014) have noted that not all citizen-consumers may have the same capabilities or conditions to make choices that maximise their own satisfaction. In this light, EU policy makers are increasingly turning their attention to the so-called "vulnerable consumers" (European Commission, 2012). Vulnerable consumers have been defined by Andreasen and Manning (1990, p. 13) as those "*at a disadvantage in exchange relationships where that disadvantage is attributable to characteristics that are largely not controllable by them*". Vulnerability can be derived from higher difficulties to obtain or assimilate the information for decision-making and/or from a higher risk for the own welfare associated to inadequate consumption decisions (Burden, 1998). As vulnerable consumers do not constitute a separate section of the population (George et al., 2011; Hogg et al., 2007), the analysis on vulnerability focuses on socioeconomic characteristics representative of risk of vulnerability or potential vulnerability (Burden, 1998; George et al., 2011; Xavier, 2008). A key factor in this regard, as explained in detail in section 3, is educational attainment, on which this study is focused. As noted by Clifton et al. (2011) and Jilke (2014), citizens' potentially vulnerable as consumers generally count on less economic, cultural and social resources for accessing and processing the information required for decision-making. As a result, they will be more likely relying on cues and biases in decision-making processes (i.e. choice overload, loss aversion, or the use of cognitive heuristics), which may move them further

away from optimal decisions. The incipient existing evidence provides support to this statement and points towards a gap between vulnerable and non-vulnerable groups of citizen-consumers in terms of the prices and quality of the services they receive. For example, in their analysis of the effects of market-oriented reforms on citizens' satisfaction for the EU-15 countries, Fiorio et al. (2007) observed that the elderly, the unemployed and those with a lower educational attainment were showing lower satisfaction levels with electricity, gas and fixed telephone services. For electricity, Fiorio and Florio (2011) found that those with lower educational attainment and the unemployed find the services they use to be less affordable. Similarly, Clifton et al. (2011) found in a study focused on Spain and the UK that citizens with lower educational attainment, the elderly and those not working were found their services to be less affordable. Based on this evidence, we expect that vulnerable citizen-consumers facing difficulties in identifying the services that best match their subjective price-expectations. We thus formulate the following hypothesis which expects a affordability-gap between vulnerable and non-vulnerable citizen-consumers:

Hypothesis 1: Citizens with socio-economic characteristics representative of potential vulnerability as consumers perceive the price of public infrastructure services to be less affordable.

As noted by Hogg et al. (2007), the increasing complexity of consumption decisions and information asymmetries can be a source of potential difficulties for disadvantaged individuals and groups in order to maximise their satisfaction in the marketplace. Jilke and Van de Walle (2013) relate this question with the debates on the introduction of choice in public services in the UK. In this debate, authors as Appleby et al. (2003), Ayeni (2000) and Needham (2007) argue that choice would mainly benefit the middle-class, whilst the poorer and less educated citizens face relatively high transaction costs associated to the process of choosing, and thus inequalities would be increased. It can be expected that in countries where the structure of the market is less concentrated, and thus competition is higher, the possibilities for choice between providers increase, but also the complexities of the decision-making processes. Increasing complexity could affect potentially vulnerable citizen-consumers in a different manner. It can be expected that, due to their lower resources for decision-making, difficulties would tend to increase more strongly for potentially vulnerable citizen-consumers than for their counterparts. As a result, their risk of experiencing cognitive biases in decision-making, moving them further away from optimal decisions, would also particularly increase. This would then widen the affordability-gap between potentially vulnerable citizen-consumers and their counterparts, leading to our second hypothesis:

Hypothesis 2a: The gap between citizens with socio-economic characteristics representative of potential vulnerability as consumers and the rest in self-perceived affordability of public infrastructure services is ***higher*** *in those countries where competition in terms of market structure is higher.*

However, it could also be argued that competition provides increasing opportunities of choice to all segments of society and that potentially vulnerable citizen-consumers would be also able to take advantage of it, compensating their hypothetical higher difficulties for decision-making. According to proponents of the choice agenda, greater provider choice would provide increasing opportunities not only to better positioned citizen-consumers who more easily can access them, but to all layers of society, thus leading to reduce inequalities among socio-economic groups (Giddens 2003; Le Grand 2005). Moreover, following Dowding and John (2009), due to the higher relative benefits (in terms of their economic capacity) that a better offer would provide to potentially vulnerable citizen-consumers, their transition costs of the searching process could be lower. We thus formulate an alternative hypothesis:

Hypothesis 2b: The gap between citizens with socio-economic characteristics representative of potential vulnerability as consumers and the rest in self-perceived affordability of public infrastructure services is ***lower*** *in those countries where competition in terms of market structure is higher.*

The arguments on the discussion of the previous hypotheses can be similarly applied to an indicator that refers to the *effective competitive functioning* of markets – in terms of citizen-consumers' actual switching rates (competition from the demand-side) – instead of the *competitive structure* of the markets (competition from the supply-side) as in hypotheses 2a and 2b. Effective competitive market functioning means that in liberalised markets citizen-consumers actually take up their chances to send market signals and thus provide effective market pressures to service providers by switching among different service offers. Indeed, this effective competitive market functioning can be regarded as a result of the competitive market structure (necessary but not sufficient condition), but is also enhanced by regulatory policies and its interaction with citizen-consumers' market behaviour (Armstrong and Sappington, 2006; European Commission, 2012). It thus reflects the demand-side of competition, justifying its separate consideration. Conforming to arguments for hypothesis 2a, it can be assumed that a greater switching behaviour in national markets reflects mainly the switching of better-positioned citizen-consumers (see Jilke, 2014). However, their vulnerable counterparts, in contrast, would face greater obstacles in these markets due to the increasing complexity

they deal with – for example by navigating through the various service offers. As a result, they would either end-up making sub-optimal decisions, or sticking with the market incumbent by avoiding switching at all, as shown by Jilke (2014). This leads to the following hypothesis which expects an increase in the affordability-gap between vulnerable and non-vulnerable consumers:

Hypothesis 3a: The gap between citizens with socio-economic characteristics representative of potential vulnerability as consumers and the rest in self-perceived affordability of public infrastructure services is ***higher*** *in those countries where competition is higher in terms of frequency of switching.*

However, it could also be argued that a higher overall frequency of provider switching in these markets does not require that vulnerable citizen-consumers switch to increase their individual utility, but that the switching behaviour of a small group of well informed citizen-consumers is sufficient to create market pressures for service providers. Studies in the area of school choice, for example, have applied such a theory of marginal consumers and exemplified that for competitive public service markets to work it is sufficient that a minority of well informed consumers (the so-called marginal consumers) shop around for schools (Buckley and Schneider, 2003; Schneider et al., 1998; Teske et al., 1993). The switching (and complaining) behaviour of these marginal consumers represent the market signals they send to service providers thereby creating market pressures to adjust services' prices and quality. In other words, from the viewpoint of a theory of marginal consumers we can assume that once a critical mass of (non-vulnerable) citizen-consumers starts switching among service providers, this creates positive externalities to all citizen-consumers, including those who are potentially vulnerable in the marketplace. Indeed, prior research has shown that non-vulnerable citizen-consumers are considerably more likely to switch public infrastructure providers (Jilke 2014). Therefore, we expect that once national switching rates go up, the affordability-gap between potentially vulnerable and non-vulnerable citizen-consumers decreases, leading to our final hypothesis:

Hypothesis 3b: The gap between citizens with socio-economic characteristics representative of potential vulnerability as consumers and the rest in self-perceived affordability of public infrastructure services is ***lower*** *in those countries where competition is higher in terms of frequency of switching.*

3. Data and Research Method

To address our research question we undertook a multilevel analysis of data combining information about subjective affordability evaluations and market structures across the EU-25 member states. Individual level data were gathered from Eurobarometer (EB) 65.3 (European Commission, 2007). EB are surveys promoted by the EC, conducted twice a year with the aim of assessing public opinion within the EU. EB 65.3 was fielded between May 5 and June 11 2006 and included all EU-25 member countries. It is representative at the national level for a population aged 15 and above (GESIS, 2012).

Dependent Variable

For citizens' self-perceived affordability of public infrastructure services we used EB 65.3 in which respondents were asked to indicate "In general, would you say the price of (fixed telephone/ electricity supply) is affordable, or not? By that, I mean that I would like to know if you are able to afford the services you need". Only those having access to the respective service were asked. Answer possibilities included (1) "Affordable, and (2) "Not affordable', including a spontaneous category "Excessive". Perceiving a service as being affordable is interpreted as a subjective evaluation of the personal utility that one derives from the infrastructure services that respondents use (see also Bacchiocchi et al., 2011; Clifton et al., 2011; Fiorio and Florio, 2009; 2011; Florio 2013). As reflected in Figure 6.1, substantial differences in mean levels of self-perceived affordability with the price of public infrastructure services exist across EU countries.

A potential difficulty when using comparative public opinion data is cross-cultural measurement (non-) equivalence (see chapter 7 for an overview). Our analysis is

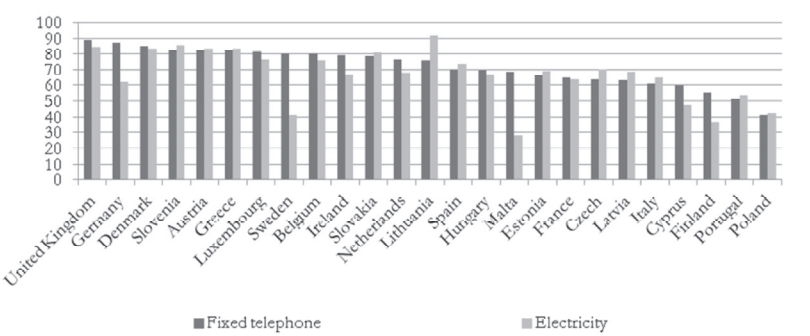

Figure 6.1: Percentage of citizen-consumers' affordability of public infrastructure services. *(Nfixed=22,212; Nelectricity=23,919).*

based on a single item dependent variable which, as encompassing only two clearly defined response categories, and given its accurate translation (including back-translation and proper handling during fieldwork, see for EB 65.3 GESIS, 2012), reduces the space for misinterpretation. Furthermore, we checked whether item non-responses for our affordability measure differed substantially across countries, and found only little evidence for it.[30]

Independent Variables

Citizen-consumers' vulnerability is not directly observable but it can be analysed through socio-economic characteristics particularly representative of it. In this study, we focus on low educational attainment, a dimension commonly used as a factor representative of consumers' potential vulnerability (Burden, 1998; George et al., 2011). Following Hogg et al. (2007), education is strongly related with the two key elements that characterise consumers' resources for participating in the market: their skills and their assets. A lower educational attainment is associated with more limited resources, and thus higher difficulties for accessing and analysing the information required for decision-making in public service markets (Adkins and Ozanne, 2005; Calero et al., 2011).. Indeed, as explained by Clifton et al. (2011) and Jilke (2014), those citizen-consumers with lower educational attainment may suffer a higher risk of experiencing biases in decision-making (such as choice-overload, loss aversion, or inertia). Illustratively, Dohmen et al. (2010) and Hjorth and Fosgerau (2011), for example, found a causal relationship between lower educational attainment, cognitive ability and subsequently individual's risk aversion. This makes educational attainment an excellent candidate for examining the relationship between citizen vulnerability and market competition for public infrastructure services. Our independent variable capturing educational attainment refers to respondents' age when finished their full-time education. It results in three different categories: basic or no education (age before 16), secondary education (age between 16 and 19) and higher education (age of 20 and over), with the last used as reference category. As we are interested in the individual within-country effects of education on self-perceived affordability and not in their structural country differences, educational level dummy-variables have been group mean centred prior to analysis (Enders and Tofighi, 2007).

According to our hypotheses, we differentiate between two dimensions of market competition. First, competition in terms of market structure (competitive market

[30] The only cases were we found substantial deviations from the mean value of "don't know" responses (5.12 for fixed telephony, and 2.22 for electricity), was within the fixed telephony sector in Finland (13.72%), Estonia (13.2%) and Latvia (16.77%). Thus, we checked for robustness by re-estimating all our models without those three countries. Result obtained did not differ from the ones presented in this study.

structure). This variable represents the degree of concentration of the market of each service in each country: that is, the level of competition understood from the supply-side. For fixed telephony, market competitive structure is measured by the Herfindahl-Hirschman Index (HHI), obtained from data of the Directorate General for Information Society of the EC (European Commission, 2008b), referred to the year 2006. HHI takes values from 0 to 10,000 (absolute concentration). For electricity, as HHI is not available for the retail markets, market competitive structure is measured by the retail market share of the 3 largest companies (the so-called concentration-index), obtained from the Directorate General for Energy and Transport of the EC (European Commission, 2008c). The information selected is referred to 2006 or the closest year available[31]. For Belgium, Netherlands and the UK, as data are not directly available, a proxy of this indicator is calculated from EUROSTAT (2008) data on the number of companies with a market share over 5% in the retail market and their cumulative market share, referred to 2006[32]. For Denmark there is no information available, and thus the analysis on this point is limited to 24 countries.

The second dimension competition is competition in terms of frequency of switching (competitive market functioning, or simply), reflecting the effective level of competition in the markets in terms of citizen-consumers' decision-making, thus competition understood from the demand side. For both services, competitive market functioning is measured by the percentage of consumers that have switched their provider in the last two years, as a country average for each service, obtained from EB 243 on switching service providers (European Commission, 2009b). Data are referred to the period 2006-2008 and is available for all the EU-25 countries. Specifically, the users of each service were asked to indicate "Have you tried to switch your [SERVICE] provider in the last two years?". From those respondents that indicated that they actually switched their service provider, we calculate the percentage of switchers per country.

Control Variables

In our analysis, we include control variables for a wide range of factors that may have an influence on our hypothesised relationships. As regards socio-economic characteristics, we control for respondents': age – including a squared term-; place of residence, differentiating between urban, semi-urban and rural areas; gender;

[31] For France, Luxembourg, Czech Republic, Poland and Slovakia, data are referred to 2007. For Italy and Spain, as the indicator is highly volatile, it is used the average of 2006 and 2007.

[32] This proxy is calculated as (Number of main suppliers / Cumulative market share of main suppliers)*3, where "main suppliers" are those providers with market share ≥ 5% (being 3 in Belgium, 5 in the Netherlands and 7 in the UK).

and the size of the household where they are currently living. EB does not contain direct information on respondent's income, or wealth status. As proxies, we control for whether respondents own a house or not, as well as for respondents' social-class, differentiating: employers, high-level non-manual workers, medium-level non-manual workers, low-level non-manual workers, self-employed in primary sector and workers. Moreover, we added two categories for those respondents who are either unemployed, or not working for another reason (e.g. students, retired persons). We also controlled for respondents' political orientation. Respondents were asked to indicate their political orientation on a scale ranging from '1' (left) to '10' (right). This type of question typically produces a lot of missing cases. Thus we grouped it into 4 categories: left (categories 1-4), middle (categories 5-6), right (categories 7-10), and one category for those who refused to answer. Finally, we control for whether the respondent was born in their country of residence, or not.

We also control for service delivery related factors. First, we control for service-usage, meaning whether respondents actually use electricity or fixed telephony services. Moreover, we control for the access difficulty to the service under consideration. Respondents were asked: "In general, would you say that access to [SERVICE] is easy or difficult for you? By that I do not mean 'affordability'". Respondents who stated that service access tends to be easy have been coded as '1', all other as '0'. We also control for how important respondents consider the service for their daily life, ranging from 'very important' to 'not at all important'. This information is available for fixed telephony, but was not asked for electricity so is not included in those models.

Besides individual level control variables, we also control for important country level characteristics. All country level variables have been mean centred prior to analysis to aid estimations and interpretability (Hox, 2002). Firstly, we control for the actual market size. For fixed telephony, we use the number of subscriptions in 2006 as provided by the International Telecommunication Union. For electricity services, we use the total electricity consumption, obtained from EUROSTAT and referred to the year 2006. Furthermore, we control for economic inequalities within countries by using the GINI coefficient of equivalent disposal income in 2006, provided by EUROSTAT. We also control for the years since each country opened the market to domestic consumers in 2006, accounting for the differences in the time when countries formally liberalised their markets. This information was obtained from EC (2005) for electricity, and from Conway and Nicoletti (2006), complemented with Bismut (2004), for fixed telephony. Finally, we control for

price dynamics in the markets by including a 2-year lagged variable (2004-2006) of consumer prices[33], in Euros per kilowatt-hours, obtained from EUROSTAT.

Method

The research design required modelling the relationship between different people's self-perceived affordability (individual level) and public service market characteristics (country level). Thus we use multilevel modelling techniques to deal with the nested structure of our data (individuals nested in countries). From a statistical point of view, using multilevel methods allows to account for potential clustering effects and unobserved heterogeneity across countries by providing random country intercepts and slopes, so as residual components at both levels of analysis, individuals and countries (Hox, 2002). Furthermore, it permits us simultaneously to estimate individual and country level effects on citizens' self-perceived affordability of public infrastructure services. Because of the binary nature of our dependent variables, we estimate a logistic multilevel model. Using multilevel techniques requires a minimum of 20 country units to produce approximately unbiased estimates for binary random intercept models (Stegmueller, 2013). Including a total of 25 countries, we are comfortably able to fulfil these minimum conditions.

4. Research Findings

Table 6.1 displays results for fixed telephony and the results for electricity are presented in Table 6.2. We display only the results for the main variables of interest, but provide full results for all predictors in Appendix B. We first assess an intercept only model for both services under consideration to determine whether multilevel modelling is actually needed, and to assess how much of the total variance in affordability can be attributed to country differences (model 0). We then estimate a model that includes all control variables for each of the two different indicators of competition (models 1 and 3). Following, we test whether we find an interaction between the degree of competition and citizen-consumers' educational attainment, as representative of their potential vulnerability (models 2 and 4). We report odds ratios and their respective standard errors in parentheses.

For fixed telephony (Table 6.1), model 0 exhibits a significant chi-square for a likelihood ratio test that compares it against an ordinary logistic regression[34]. It rejects

[33] Due to limited data availability for a significant number of countries, it is not possible to include information on prices at a longer term.

[34] Likelihood ratio test vs. logistic regression yielded a chi-square difference of 1468.49 and was highly significant (99% level).

Table 6.1: Results fixed telephony.

FIXED TELEPHONY					
	Model 0 Intercept only	Model 1 Market competitive structure	Model 2 Market competitive structure + interaction	Model 3 Market competitive functioning	Model 4 Market competitive functioning + interaction
Intercept	2.779** (.337)	2.356** (.501)	2.393** (.509)	2.354** (.501)	2.456** (.524)
Market comp. structure (reversed)		1.000 (.000)	1.000 (.000)		
Market comp. Functioning				1.002 (.014)	1.002 (.014)
Basic education (Ref: higher education)		.794** (.043)	.794** (.043)	.794** (.043)	.780** (.042)
Secondary education (Ref: Higher education)		.872** (.038)	.872** (.038)	.872** (.038)	.872** (.038)
Mark. comp. structure X Basic educ.			1.000* (.000)		
Mark. comp. structure X Second. educ.			1.000 (.000)		
Mark. comp. functioning X Basic educ.					1.017** (.006)
Mark. comp. functioning X Second. educ.					1.002 (.005)
Variance: country intercept	.362 (.104)	.293 (.085)	.293 (.085)	.295 (.085)	.294 (.085)
Log likelihood	−11,857.64	−11,064.80	11,062.52	−11,064.85	−11,059.83
Interclass correlation	.099	.082	.082	.082	.082
N (individuals)			21,701		
N (countries)			25		

Odds ratios are reported with SE in parenthesis.
*** sig. at 1%; * sig. at 5%.*

the null that a conventional logit model is performing better than a multilevel model. Furthermore, the interclass correlation shows that almost 10% of the total variation in affordability of fixed telephony services can be attributed to country differences. In model 1, we examine the effects of the competitive market structure and the effects of individuals' educational status on self-perceived affordability. We find that the competitive market structure has no direct effect on affordability evaluations. Whereas for education, in line with hypothesis 1, the odds of less educated individuals on perceptions of affordability of the price are significantly smaller when compared to those that completed higher education; the same holds for those

respondents with secondary education, however, to a slightly lesser extent. In a next step, we assess the interaction effect between individuals' educational attainment and competitive market structure. While the interaction terms turn statistically significant, their inclusion do not improve the model's fit[35]. We thus conclude that the observed interaction effect is not of substantial magnitude.

Within model 3, we assess the effect of competitive market functioning (percentage of switchers per country) on citizens' self-perceived affordability of fixed telephony services. Analogously as in model 1, there is not a direct effect of competitive market functioning on affordability evaluations. Looking at model 4, we assess whether there is an interaction between educational attainment and competitive market functioning. Here, the interaction term yields a p-value statistically significant at the 99% level, and its inclusion significantly improves model fit[36]. It shows that in those countries where competitive market functioning is higher (measured by the percentage of switchers), the gap in self-perceived affordability between the less-educated and their higher educated counterparts decreases.

The identified interaction between educational attainment and competitive market functioning requires further inspection of its effect size and direction. Figure 6.2 displays the marginal effects of being low educated on people's affordability evaluations, contingent on the degree of competitive market functioning in terms of percentage of switchers (holding all other variables constant at their mean values). Figure 6.2 shows that in markets where a higher percentage of citizen-consumers switch on average, the gap between the less-educated and their higher educated counterparts decreases. Even, considering a 95% confidence interval, this gap is wiped out at a threshold of 22% switchers.

Table 6.2 summarises the empirical results for the electricity sector. Comparing our null model with a conventional single level logit, the likelihood ratio test reflects that using a hierarchical model is a better choice[37]. Furthermore, the model's interclass correlation indicates that 16% of the total variation of people's affordability evaluations can be attributed to country differences. Model 1 reflects that competitive market structure has no effect on affordability. Also, model 1 shows that the less-educated are less likely perceive the prices of their electricity services as

[35] Performing a likelihood ratio test for two additional degrees of freedom and a chi-square difference of 4.56 yielded no statistically significant improvement.

[36] A chi-square difference of 10.03 with two degrees of freedom improved our models fit at the 99% significance level.

[37] Likelihood ratio test vs. logistic regression yielded a chi-square difference of2517.64 and was highly significant (99% level).

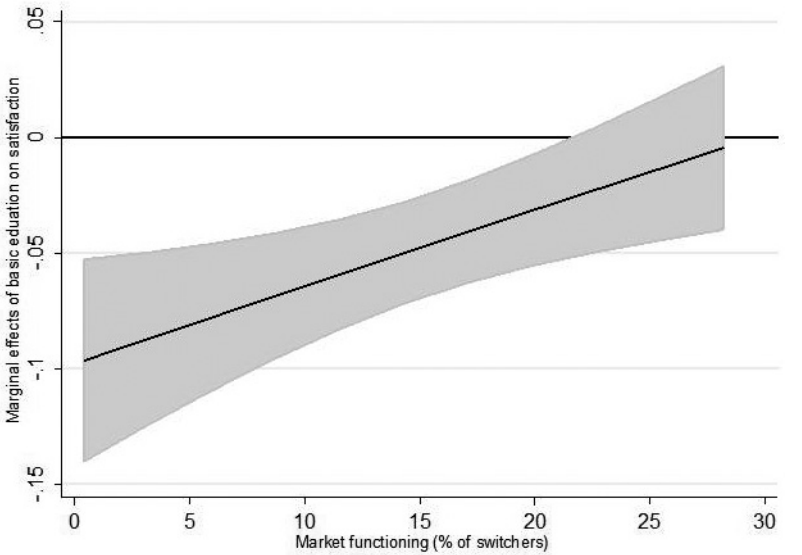

Figure 6.2: Marginal effects of being low educated on self-perceived affordability contingent on market competitive functioning (95% confidence intervals) in the fixed telephony sector.

affordable, when compared to those with a higher level of education, in accordance to our initial assumption. In a next step, model 2 reflects that the interaction terms between educational attainment and competitive market structure are not statistically significant, nor do they increase the model's fit. From model 3, we obtain the result that competitive market functioning (percentage of switching), in turn, has also no direct effect on individual's self-perceived affordability in the electricity market. Model 4 assesses the interaction between competitive market functioning and educational attainment. Like for the fixed telephony (Table 6.1) analysis, this interaction term is statistically significant, and it significantly improves the model's fit[38]. It reflects that the affordability-gap between the less-educated and their higher educated counterparts decrease as the percentage of switchers in the market increases.

We further evaluate the cross-level interaction between educational attainment and competitive market functioning by calculating the marginal effects of being low educated on affordability evaluations contingent on the percentage of switchers in the market. Figure 6.3 displays the results of these calculations, keeping all independent variables constant at their mean values. As observed, as the percentage of

[38] A chi-square difference of 9.79 with two additional degrees of freedom was significant at 99%.

Table 6.2: Results electricity.

ELECTRICITY					
	Model 0 Intercept only	Model 1 Market competitive structure	Model 2 Market competitive structure + interaction	Model 3 Market competitive functioning	Model 4 Market competitive functioning + interaction
Intercept	2.230** (.359)	2.198** (.497)	2.191** (.496)	2.292** (.505)	
Market comp. structure (reversed)		.992 (.007)	.992 (.007)		
Market comp. functioning				1.055 (.031)	1.055 (.031)
Basic education (Ref: higher education)		.830** (.041)	.829** (.041)	.827** (.041)	.829** (.041)
Secondary education (Ref: Higher education)		.869** (.034)	.868** (.034)	.866** (.034)	.866** (.034)
Mark. comp. structure X Basic educ.			1.002 (.002)		
Mark. comp. structure X Second. educ.			1.000 (.002)		
Mark. comp. functioning X Basic educ.					1.011* (.005)
Mark. comp. functioning X Second. educ.					.996 (.004)
Variance: country intercept	.641 (.184)	.350 (.103)	.350 (.103)	.319 (.092)	.319 (.093)
Log likelihood	−13,327.63	−12,682.10	12,681.65	−13,078.75	−13,073.85
Interclass correlation	.163	.096	.096	.088	.089
N (individuals)	23,352	22,416		23,352	
N (countries)	25	24		25	

Odds ratios are reported with SE in parenthesis.
*** sig. at 1%, * sig. at 5%.*

switchers in a country increases, the affordability-gap between the less-educated and their higher educated counterparts decreases. Furthermore, considering a 95% confidence interval, this gap is eventually wiped out once the percentage of switchers exceeds 13%.

Summing up these empirical results, a gap in the self-perceived affordability of electricity and fixed telephony services is observed between less-educated citizen-consumers and their better educated counterparts, even after controlling for a wide range of socio-economic and country-level variables. This is in line with previous

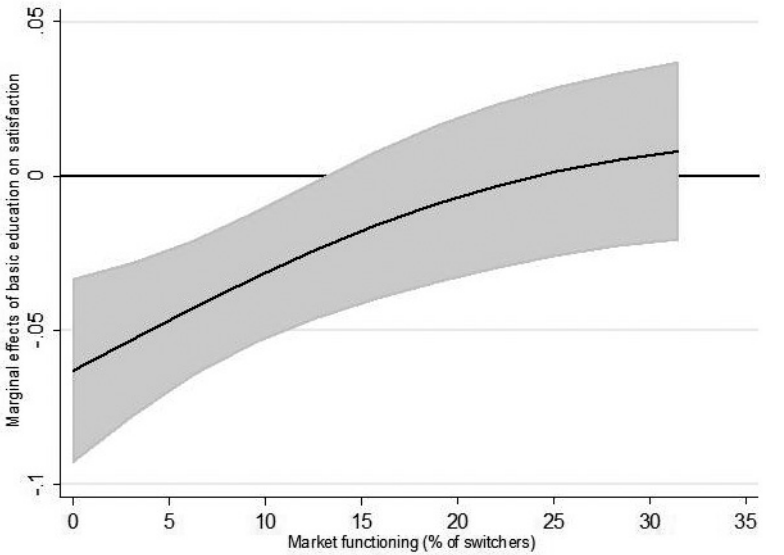

Figure 6.3: Marginal effects of being low educated on self-perceived affordability contingent on market competitive functioning (95% confidence intervals) in the electricity sector.

studies. We do not find an effect of market's competitive structure on this gap, however, for both services, market's competitive functioning (measured by the percentage of switchers) is associated with a significantly lower affordability-gap.

5. Discussion and Conclusions

The findings lead us to confirm the affordability-gap hypothesis for both service sectors: lower educational attainment, representing potential vulnerability, is indeed significantly related with lower affordability evaluations of both services. Or in other words, lower educated citizen-consumers perceive on average the affordability of their services as lower. However, both hypotheses 2a and 2b are rejected: the competitive market structure in a country alone does not show any significant or substantial effects on the gap in affordability. However, whilst hypothesis 3a is also rejected, evidence supporting hypothesis 3b is obtained: in both service sectors, the affordability-gap tends to be lower as the frequency of switching (competitive market functioning) is higher in a country. Eventually the gap is wiped-put once there are more than 22% or 13% switchers in the telephony and electricity sector, respectively. Thus whilst the market competitive structure (supply-side) does not lead to any change in the observed affordability-gap in the countries under

scrutiny, the competitive functioning of these markets – this means when national markets exhibit higher switching rates (demand-side) – does substantially reduce the gap in these countries.

Our findings suggest that liberalisation and market competitive structures alone are not the principal factor leading to reduce the gap between potentially vulnerable citizen-consumers and the rest in terms. Instead, markets able to significantly extend switching can be successful in closing this gap. Thus it seems that it is the demand and not the supply side of competition that is crucial here. Once there is a critical mass of switchers, the positive externalities of the market pressures they create spread among citizen-consumers in these markets – including those citizens who are potentially vulnerable as consumers. Furthermore, these results may suggest that the value of choice may come not from an increasing number of alternatives or simply competition per se, but from providing information, advice and help for citizen-consumers' active behaviour in the market. Thus, the evidence obtained in this study supports the increasing orientation of EU regulatory policies towards the demand-side, by focusing on empowering consumers in the market (European Commission, 2012), after having focused in the market-oriented reform of public infrastructure services from the supply-side during previous decades.

Finally, the characteristics of each service could be another significant factor to be considered. In electricity, although the EU countries have generally introduced market-oriented reforms, only few of them have achieved a significant frequency of switching. Whereas telecommunications have experienced a deep technological change, with the emergence of services substitutive of fixed telephony, as mobile telephony and internet. The analysis by Jilke (2014) has obtained that as the number of providers increases, the gap as regards education in the probability of switching does not experience significant changes for fixed telephony, but increases for mobile telephony. The degree of technological innovation, variability and complexity can make the difference, as requiring higher capacities for decision-making, but also the existence of habits and previous experience in consumption, enabling to turn to a "satisfaction option". Therefore, the characteristics of each market matter, and thus it would not be always possible to generalise the results obtained in a particular service, neither on a particular country, to the rest. Thus, this study opens considerable space for further research on this topic, aim at improving the understanding of consumers' attitudes and motivations in these markets increasingly complex, but still essential for their daily life and for EU policy action, as public infrastructure services. A better knowledge of the consumers' perspective is required if EU regulatory policies aim to continue improving the functioning of these markets not just from the supply-side, but also from the demand-side.

CHAPTER 7

We Need to Compare, but How? Cross-national Measurement Equivalence in Comparative Public Administration

This chapter is an extended version of Jilke, Meuleman and Van de Walle (2014).

1. Introduction

Consider the following survey item: *"Overall, how satisfied are you with your electricity supplier? Please give me score from 0 to 10 where 0) means that you are not satisfied at all, and 10) that you are fully satisfied"*. This is one out of a battery of items that taps citizens' satisfaction with public services across a wide range of countries. The underlying assumption of asking the same set of items to respondents in different national populations is that their answers are supposed to be comparable. In other words, it is assumed that perceptions of what satisfaction means, and the way in which people use assigned scales are equivalent across countries, allowing for meaningful comparisons. But is the general notion of what a satisfactory public service is really equivalent across countries, regions, (groups of) individuals or even over time? And are patterns of response styles the same across different cultures? In this article, we introduce two major techniques for detecting, and correcting non-equivalence into the field of public administration. By means of concrete examples, we show how these methods can be implemented in applied research.

Comparing public administrations, public managers or citizens interactions with, and attitudes towards government across countries is gaining ground in public administration research (e.g. Jilke, 2014; Kim et al. 2012; Pollitt and Bouckaert, 2011; Van Ryzin, 2011). This is accompanied by an increase in availability of cross-national surveys which contain questions relevant for public administration research, such as the *International Social Survey Programme*, the *Eurobarometer*, the *COCOPS survey* of public managers, or the *COBRA survey* of government agencies' executives, among many others. Making use of such cross-national survey data gives us the opportunity to test the geographical range of social theories by assessing them in many different contexts. For example, one may study whether the (dis)confirmation effect of individuals' expectations on their satisfaction with services (e.g. James, 2009; Morgeson, 2013; Van Ryzin, 2006, 2013) is of the same magnitude within countries which are embedded in different national systems of service delivery. Moreover, having survey data from numerous countries enables us to investigate various micro-macro relations by utilizing data from the individual *and* the country level. For example, one may look at the potential moderating effects of public spending, or the political context more broadly, on the positive relationship between satisfaction with public services and trust in government. Such cross-level interactions permit us to more closely look at interesting relationships between context and individuals, allowing us to explicitly test contextual theories (see for example O'Toole and Meier, 2013).

However, when respondents in different countries regard measurement constructs in a different manner, or exhibit culturally influenced response patterns, we

typically obtain biased survey measures (Poortinga, 1989; Van de Vijver and Leung, 1997). Practically speaking, the response of a person in country A, say to the item on satisfaction we used as an example, may have the same scale-position than the response of another person in country B, but it could mean something entirely different if the way respondents interpret or respond to it differs substantially. By simply looking at mean levels of survey responses, however, we do not know whether the answers of both respondents can be meaningfully compared. This puts empirical tests at risk as we cannot confidently claim measurement equivalence, and may end up comparing apples and oranges. In such a case, results from statistical estimations, so as the theoretical implications that we draw from cross-national data, are invalid and can lead to spurious conclusions (Davidov et al., 2014).

Responding to recent calls to cross-nationally assess public administration theories of context (O'Toole and Meier, 2013), in this article we provide an examination of the concept of cross-national measurement equivalence in public administration, and how to proceed in establishing the comparability of survey measures. It is structured as follows: first we introduce the concept of measurement equivalence and elaborate on the importance of utilizing appropriate techniques to deal with measurement non-equivalence in comparative public administration. We report from a systematic literature review of empirical studies using cross-national surveys in public administration and investigate if, and how those works haven taken the issue of measurement (non-)equivalence into account. Consecutively, we introduce two procedures on how to detect, account and even explicitly correct for measurement non-equivalence, namely multiple group confirmatory factor analysis (MGCFA) and item response theory (IRT). While MGCFA, is most appropriate for continuous data, IRT modelling is best suited for ordered-categorical (or binary) items[39]. We, furthermore, illustrate the application of these statistical procedures using two empirical examples on 1) citizens' satisfaction with public services, and 2) trust in public institutions. Our findings indicate how appropriately dealing with non-equivalence accounts for different forms of biases, which might otherwise stay undetected. We conclude our article by discussing the implications for cross-national survey research within the discipline. In doing so, this article ties in with the recent movement in making quantitative research in public administration more rigorous and innovative (e.g. Konisky and Reenock, 2012; Lee, Benoit-Bryan and Johnson, 2011; Meier and O'Toole, 2012; Zhu, 2012; see also Gill and Meier, 2000; DeLorenzo, 2001, Wright, Manigault and Black., 2004), and thus contributes to the methodological advancement in studying public administration beyond domestic borders.

[39] MGCFA can also be applied to test for measurement equivalence with binary/ordinal items via suitable estimators (see Millsap and Yun-Tein, 2004).

2. Measurement Equivalence in Comparative Public Administration

In order to expand public administration theories to other cultural settings, researchers often have to rely on secondary data. Thus they have little or no control over survey design procedures that would help them to establish the cross-national equivalence of their items, for example through the use of anchoring vignettes (King et al., 2004). In such a case, scholars aiming to utilize cross-national survey data have to find appropriate ways to make sure that their measurement constructs, such as trust, satisfaction or motivation, to name a few, are equivalent across countries. If this is not done, cross-national comparisons are likely to be invalid (cf. Vandenberg and Lance, 2000). Thus measurement non-equivalence can be considered as a serious threat to comparative public administration survey research.

In recent years there has been a growing awareness in applying post-survey techniques to assess measurement (non-)equivalence. In line with this, several statistical methods have been applied for testing, including MGCFA, and IRT. This development can be observed across a wide array of disciplines within the social sciences. However, within public administration research this seems largely ignored. To illustrate this point, we conducted a systematic literature review of journal articles in public administration that make use of cross-national survey data for the time period 2001 till 2012. The following Social Science Citation Index listed journals were consulted: *Administration & Society*, *American Review of Public Administration*, *International Review of Administrative Sciences*, *Public Administration*, *Public Administration Review* and *Journal of Public Administration Research and Theory*.[40] The review resulted in a total of 19 articles, with almost 75 per cent (14 articles in total) of the studies being published since 2008 – emphasising the growing interest in cross-national survey research in the discipline. Following, all articles were reviewed with regard to 1) acknowledging the possibility of measurement non-equivalence for the used data, and 2) whether authors have taken any measures to test for non-equivalence, and/or corrected for it. It was found that only two articles from our review mentioned the possibility of cross-national non-equivalence of their survey items. From those two articles, only one goes the next step in testing for non-equivalence by means of a MGCFA. These results are worrisome given the share of studies that have been produced without appropriately dealing with the possible non-equivalence of their survey measures. It, indeed, suggests the limited awareness of public administration scholars about applying post-survey techniques

[40] With the exception of the *International Review of Administrative Sciences*, which we added because of its explicit comparative scope, these journals have been included by previous reviews on research methodology in public administration, because they are thought to be the mainstream journals within the discipline (see for example Brower, Abolafia and Carr, 2000; Lee, Benoit-Bryan and Johnson, 2012; Wright, Manigault and Black, 2004).

to deal with the possibility of measurement non-equivalence (see also Kim et al., 2012), and highlights the importance of an accessible primer on measurement equivalence in comparative public administration.

3. A Conceptual Framework Linking Measurement Bias with Equivalence

Measurement equivalence refers to a certain aspect of the validity of survey items that tap into an underlying latent concept, such as 'satisfaction'. It means that *"[…] under different conditions of observing and studying phenomena, measurement operations yield measures of the same attribute*" (Horn and McArdle, 1992: 117). Following this definition, for measurement constructs to be equivalent, two attributes must be met. First, the unobserved latent trait must share the same meaning across different groups. For example, this is met when citizens' notion of what satisfaction with services means to them is equivalent across countries. Second, the examined latent concept needs to be scaled equally across countries – meaning that it is measured using the same metric. If this does not hold, one may speak of a measurement construct that exhibits no equivalence across groups, or in our case, countries (Meuleman and Billiet, 2012).

Measurement non-equivalence can stem from a variety of different sources, with all of them being related to different aspects of biases. Conceptually, three major types are distinguished: 1) construct bias, 2) method bias, and 3) item bias (cf. Van de Vijver and Leung, 1997; Van de Vijver, 2003). Construct bias refers to the dissimilarity of latent concepts across countries. It means that the configuration and interpretation of a hypothetical construct, such as 'satisfaction' or 'trust', may not be shared among different countries. In such a case, latent concepts cannot be easily generalised to other cultural settings. One interesting example of construct bias would be the meaning of the well-known political left-right continuum, which Piurko and colleagues (2011) found to be different across Western and Eastern European countries (cf. Davidov et al., 2014). This comes from the fact that 'left' and 'right' are associated with different aspects of political ideologies, including liberalism, traditionalism, and post-communism. While 'left' and 'right' have coherent meanings in liberal and traditional countries (such as Sweden, or Greece), this political dichotomy has a different meaning in post-communist countries: "*[...]"left" is sometimes linked to Western liberalism and sometimes to communism in these countries and these links may differ across countries*" (Piurko et al. 2011: 555). Matching up the answers of the commonly used political left-right scale across those countries, the authors conclude, would thus be like comparing apples and oranges.

The second type of bias, method bias, refers to all types of biases that come from methodological procedural aspects of a survey. They include i) the incomparability of national samples, for example by using different national sampling schemes, ii) cross-cultural differences in response behaviour, and iii) systematic differences across countries in the survey communication between interviewer and interviewee. An example in this regard would be extreme response style behaviour where respondents from certain cultures have the tendency to select the end point of a given item scale (Johnson et al., 2005). In such cases, respondents across countries may share the same scale position, but not the same meaning attached to it. This could potentially lead to a shift in the average mean score suggesting country differences which are only an artefact of these method effects.

The third type of bias is called item bias, or differential item functioning. It basically means that different people understand or interpret the very same survey item in a different way. This kind of bias directly relates to disfunctioning at the item level. An item is said to be biased "*[…] if respondents with the same standing on the underlying construct (e.g. they are equally intelligent), but who come from different cultures, do not have the same mean score on the item*" (Van de Vijver, 2003: 148). Common sources of item bias are poor translations and/or ambiguous items, cultural differences in the connotative meaning of item content, or the influence of cultural specific nuisance factors such as the involvement of social desirable answering behaviour towards specific items.

These types of biases are linked to different forms of measurement non-equivalence. In order to relate bias with measurement non-equivalence, we draw upon the generalised latent variable framework (Skrondal and Rabe-Hesketh, 2004). Here, it is commonly assumed that theoretical concepts (latent traits), such as trust or satisfaction, are not directly observable, but are inferred from multiple observed manifestations of the latent trait (Bollen, 2002; Davidov et al., 2014). Consider satisfaction with electricity services as our latent trait. Since this concept cannot directly be observed, one may use a set of different survey items that tap into citizens' levels of satisfaction. Examples would be items on their satisfaction with the price, the service quality, or any other aspects that taps into satisfaction with electricity services. Say we would measure citizen satisfaction with electricity services using multiple items across various countries, we can now test for measurement equivalence across those countries by "*[...] comparing empirical relations between the latent variable and the indicators across populations. Similarity of these relationships (as reflected by the measurement parameters) is taken as evidence supporting the hypothesis of measurement equivalence*" (Davidov et al., 2014: 19).

Using the generalised latent variable framework, cross-national researchers typically distinguish between three types of non-equivalence: configural, metric and scalar equivalence (Steenkamp and Baumgartner, 1998; Vandenberg and Lance, 2000). These types are hierarchically ordered, meaning that lower levels serve as a prerequisite to establish the next higher level of equivalence. The first level of equivalence, configural equivalence, means that a measurement model exhibits the same factorial structure across all groups under investigation. In other words, it has an equivalent configuration across countries. Configural equivalence is affected by the presence of construct bias. Moreover, it is considered as the lowest level of equivalence and serves as a prerequisite for establishing metric and scalar equivalence. Thus by solely establishing configural equivalence, scholars cannot proceed with comparing groups. This can be seen from Figure 7.1, were we depict on the x-axis the measured score of a variable, and on the y-axis the latent score of the associated latent trait for two groups (e.g. respondents in two different countries) that exhibit configural equivalence only. We can see that comparisons across groups are not possible since a one unit increase in group A has a much stronger magnitude than in group B. Thus it does not permit comparing regression coefficients across groups. Moreover, both groups have different scale origins. Hence we also cannot compare latent group means because the position on the observed items across groups is not equally corresponding with the associated score for the latent trait.

The next level of measurement equivalence is metric equivalence. It assumes that the scale intervals, or metrics, that measure the latent construct are equal across coun-

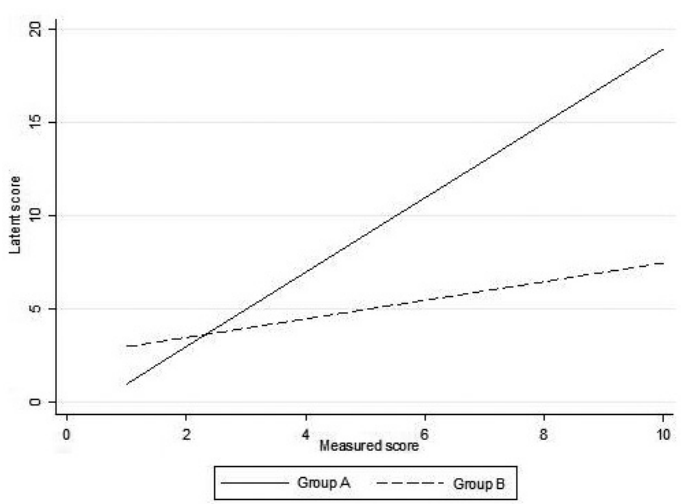

Figure 7.1: Configural equivalence.

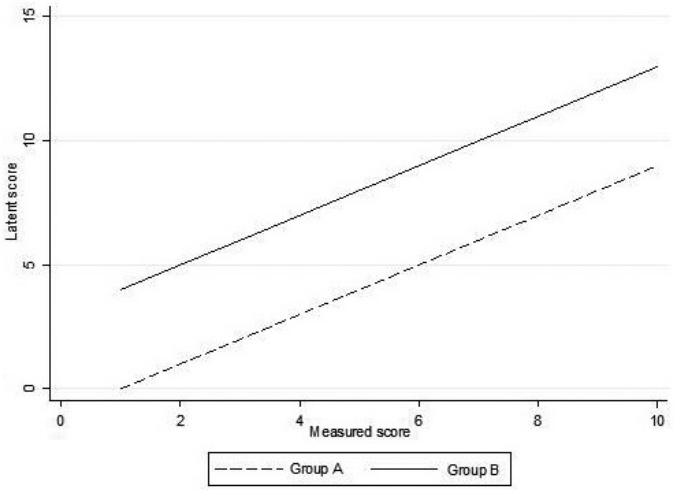

Figure 7.2: Metric equivalence.

tries. As a consequence, a one unit increase on a scale that exhibits metric equivalence has the same meaning across groups. It is affected by method and item bias. Figure 7.2 exemplifies a hypothetical latent construct that exhibits metric equivalence graphically, using simulated data. While the scale interval is equivalent across groups, meaning they can be meaningful compared, both slopes still have different origins. Thus metric equivalence permits group comparisons of regression coefficients and covariances, but not of latent means (cf. Steenkamp and Baumgartner, 1998).

The next form of equivalence, scalar equivalence, suggests that the latent variable has in addition of being measured using the same metric, the same scale origin across countries. Scalar equivalence is required when one needs to compare means across different units (cf. Meredith, 1993). This type of equivalence refers to the equality of intercepts across groups and is affected by method and item bias. If scalar equivalence holds, it shows that respondents across groups not only share the same scale metrics, but also the same scale origin. This means that they have the same score on the latent *and* on the observed variables. It can be illustrated by looking at Figure 7.3, which now depicts an identical line for both groups – note that the steepness of the slopes can vary. Practically this means that we can now compare regression coefficients and covariances, so as latent means across groups, which allows us to conduct substantial cross-national analyses.

We can conclude that higher levels of equivalence are more difficult to establish. Thus, public administration scholars who wish to meaningful compare responses

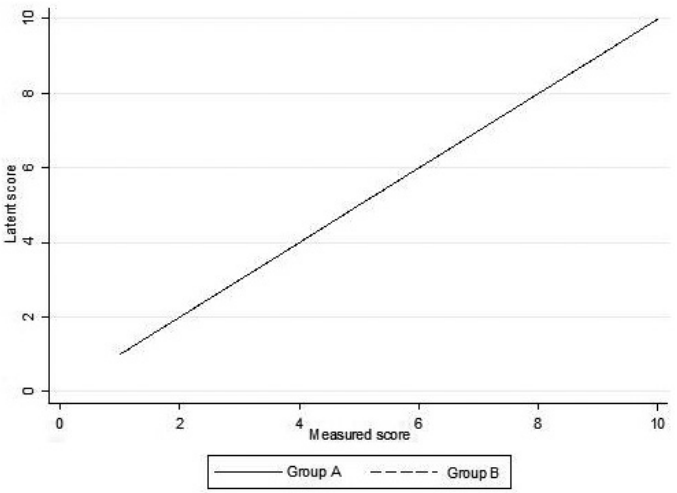

Figure 7.3: Scalar equivalence.

from cross-national surveys have to take these potential pitfalls into account by using appropriate post-survey techniques.

4. How to Detect and Deal with Measurement Non-Equivalence?

Operationalising the concept of measurement equivalence, in the following we introduce two techniques into the field of public administration of how to detect and deal with measurement non-equivalence in comparative research: 1) multiple-group confirmatory factor analysis and 2) multilevel mixture item response theory modelling. Both techniques represent two different, though related approaches towards measurement equivalence since both conceptualise measurement (non-) equivalence under the generalised latent variable framework. But while MGCFA is most appropriate for continuous data[41], IRT is specifically designed to deal with data that is of ordered-categorical nature. A distinctive feature of Confirmatory Factor Analysis (CFA) is that it can be easily employed to identify multiple latent traits that underlie the data. However, *multidimensional* IRT can also be used to effectively identify multiple latent constructs. For a standard IRT model that would be straight-forward, however, the introduced multilevel mixture IRT approach with item bias effects would become overly complex, and much more likely to converge to a local maxima. While this would be an interesting application, it is

[41] However, there exist also MGCFA estimators that allow for using items that are ordered-categorical, or binary (see for example Milsap and Yun-Tein, 2004).

beyond the scope of this article. Here, the application of a Bayesian estimator would be advisable in the case someone wants to apply the described IRT approach to a multidimensional specification (see De Jong and Steenkamp, 2010).

In the past, both approaches have enjoyed wide popularity when it comes to testing for measurement equivalence. While according to Kankaras, Vermunt and Moors (2011) differences between both techniques lie mainly in the terminology, model assumptions and procedures in testing for measurement equivalence, they also share a great deal of conceptual similarities, since both can be easily summarised within a generalised latent variable framework (Skrondal and Rabe-Hesketh, 2004). In this regard, MGCFA primarily aims at testing the equivalence of individual items and subsequently establishes different levels of measurement equivalence, including non-equivalence and partial equivalence, in an iterative process[42]. The multilevel mixture IRT model with item bias effects that is applied in the later part of the study, in turn, tests and corrects for measurement non-equivalence within a single model. Both models can be easily extended to include covariates (see Stegmueller, 2011 for an application of IRT to citizen's attitudes towards redistribution; see Davidov et al., 2008 for an application of MGCFA to study the effect of human values on anti-immigration attitudes).

5. Multigroup Confirmatory Factor Analysis

The standard, single group, CFA is designed to test a measurement model, where observed responses to a set of items are denoted as χ_i (where i= 1,…,I), and are written as linear functions of the latent construct ξ (for example 'satisfaction') they measure. The model typically also includes an intercept τ_i and an error term δ_i for each item, which can be written as follows:

$$\chi_i = \tau_i + \lambda_i \xi + \delta_i. \tag{1}$$

In equation (1), λ_i refers to the slopes, or the factor loadings, of the latent construct ξ. It denotes the change in χ_i for a one unit increase in ξ. Or in other words, it displays the regression coefficients for single items on the unobserved construct that we measure. In turn, the intercepts τ_i indicate the expected values for the observed items when the latent trait is equal to zero (cf. Steenkamp and Baumgartner, 1998).

[42] For a more technical comparison between both techniques, we refer to Kankaras, Vermunt and Moors (2011), Raju, Lafitte, Byrne (2002); and Reise, Widaman and Pugh (1993).

The described factor analytical model has been extended by Jöreskog (1971) to a multi-group setting. In this MGCFA, the same factor structure is specified for each group (i.e. country) k (where c=1,…,K) simultaneously, yielding an overall model fit. Thus we get

$$\chi^k = \tau^k + \Lambda^k \xi^k + \delta^k, \tag{2}$$

where Λ^k stands for a matrix of factor loadings, meaning it contains one value for each combination of items and the latent construct for every country. The remaining letters are vectors containing the same values like in equation (1), but with one single parameter for each group unit. Within such a framework, we can assess measurement equivalence by comparing parameter estimates across different countries. In our empirical examples, the groups are inhabitants of different countries, but one may also think of comparing different sub-national, socio-educational or professional groups, or even looking at the same groups of respondents over time. Regard the needed samples size required to perform a CFA, Kline (2013: 179-180) recommends a 20:1 respondents-parameter ratio of at least 20 respondents per each model parameter (see also Jackson, 2003), with the overall sample size preferred to exceed N=200. In the context of a MGCFA that would mean that researchers would need at least 20 respondents per parameter, per group. But in cases were no maximum likelihood estimators are employed, or items are non-normally distributed, much larger samples are needed.

Assessing different forms of measurement equivalence

While establishing measurement equivalence is an important prerequisite to meaningful compare regression coefficients, and latent factor means across groups, there exists a hierarchy of different forms of equivalence[43]. As we have mentioned earlier, it is typically differentiated between three major forms: configural, metric and scalar equivalence (Steenkamp and Baumgartner, 1998). Following an iterative process in testing for the different forms of measurement equivalence, Meuleman and Billiet (2012) propose a bottom-up strategy (see also Steenkamp and Baumgartner, 1998). This means to start with the lowest level of equivalence, that is the configural model, and then stepwise test the next hierarchical levels, first metric, and then scalar equivalence.

[43] In the CFA literature most authors use the term measurement invariance, instead of measurement equivalence. However, to remain consistent across introduced techniques and applications, we use the term equivalence interchangeable with invariance, meaning that our observed items and their factorial structures are not varying across countries and are thus equivalent.

Practically speaking, configural equivalence means that a measurement model exhibits the same patterns of salient and nonsalient factor loadings[44] across groups (cf. Horn and McArdle, 1992). It can be assessed by running an exploratory factor analysis (EFA) for each country separately and subsequently comparing the number of factors where items loaded on, so as their parameter estimates. Furthermore, one may estimate a MGCFA without constraints across groups and check whether fit indices are within an acceptable range. If configural equivalence has been established, on this basis, full metric equivalence is tested by constraining the factor loadings in the measurement model to be equal across groups. Formally, this would mean that:

$$\Lambda^1 = \Lambda^2 = \ldots = \Lambda^k. \tag{3}$$

Thus metric equivalence can be assessed by comparing multiple measurement models with constrained and unconstrained factor loadings across groups. Moreover, by determining which items' slopes are not equivalent across countries, scholars are put in the position of being able to identify non-equivalent survey items.

The lower levels of equivalence, configural and metric, serve as a prerequisite to establish the next, even stronger, level of equivalence: scalar equivalence. It is tested by additionally constraining all intercepts to be equal across countries (cf. Meredith, 1993), and can be written as follows:

$$\tau^1 = \tau^2 = \ldots = \tau^k. \tag{4}$$

However, the described forms of equivalence may not always hold to full extent. If this is the case, Byrne, Shavelson and Muthén (1989) have proposed the concept of partial equivalence. Basically, partial equivalence requires that at least two parameters per country are equivalent, while others are free to vary. In other words, as long as we have two items with invariant slopes across countries, we can establish partial metric equivalence. Moreover, if we find two items with equivalent slopes *and* intercepts, we can establish partial scalar equivalence. The basic idea behind this approach is that we need one item, the referent, to identify the scale of the latent variable, and one item to determine the metric of the used scale. In practice, this would mean that we can release invariant parameters for some items, as long as we have two calibrating items left which are equivalent across units (see also Steenkamp and Baumgartner, 1989).

[44] This does not mean that the strength of factor loadings are not allowed to differ, since there are no restrictions for their magnitude (cf. Steenkamp and Baumgartner, 1998: 80).

Determining a significant and substantial change in model fit

When testing for different levels of measurement equivalence the evaluation of model fit is of particular interest for researchers who want to determine whether releasing (or constraining) one additional parameter substantially changes model fit. The evaluation of model fit is typically based on the chi-square test (Kline, 2010). In larger samples (more than 300 respondents) chi-square is known to perform overly sensitive, however, meaning that it reaches statistical significance also for very trivial model changes (Kline, 2011: 201). Thus various authors have recommended to use alternative goodness of fit measures, such as the Root Mean Square Error of Approximation (RMSEA), or the Comparative Fit Index (CFI), among many others (Chen, 2007; Williams, Vandenberg and Edwards, 2009). However, while those alternative fit measures do not possess the same problems of sensitivity to large sample sizes as chi-square does, they have another problem that is that they do not have known sampling distributions. This makes it extremely difficult to determine an acceptable cut-off value for a statistically significant change in model fit when evaluating equivalence hypotheses (cf. Meuleman, 2012). Moreover, simulation studies have produced very different results when it comes to establishing such cut-off values. For example, Chen (2007) determined cut-off points for global fit indices. However, in a more recent simulation study Hox and colleagues (2012: 95) conclude that the "*[…] reliance on global fit indices is misleading when measurement equivalence is tested*" (see also Saris, Satorra and Van der Veld, 2009 for similar conclusions).

In line with various other authors Hox and colleagues (ibd.) recommend using more specific indicators of lack of fit, such as expected parameter changes in combination with their respective modification indices (Saris, Satorra and Sörbom, 1987; Saris, Satorra and Van der Veld, 2009; Steenkamp and Baumgartner, 1998; Meuleman, 2012; Whitaker, 2012; see also Oberski, 2014). By this, researchers would not only avoid over-fitting, and a rather data driven approach, but also be put in the position to determine a statistically significant *and* substantial change in model fit. In line with this reasoning, Meuleman and Billiet (2012) recommend using the following procedure to determine a significant and substantial improvement (or deterioration) of fit when assessing measurement equivalence: First, one needs to determine the slope (or intercept) with the highest modification index (MI) score – which reports the change in χ^2 when freeing the respective parameter. If this MI is strongly significant[45], *and* the associated standardised (STDYX) expected parameter change is of substantive magnitude, the respective parameter will be relaxed.

[45] This implies a Bonferroni-type correction to account for the fact that multiple tests are actually conducted at the same time – meaning 1 test per parameter, per country (Meuleman and Billiet, 2012; see also Saris, Satorra and Sörbom, 1987). Thus the alpha level may be varied in accordance to the number of used items and country groups.

Item response theory multilevel mixture model with item bias effects

While the use of MGCFA to detect measurement non-equivalence is often perceived as the predominant approach in cross-national research, modern item response theory (IRT) modelling offers similar advantages, with the particular difference that IRT techniques are specifically developed to deal with items that are discrete or ordered-categorical, instead of continuous. For ordered categorical items, such as Likert scales, this is the so-called graded response model (Samejima, 1969). It models items' C – 1 thresholds (where c is the item category with c=1,…,C) which are transformed on a continuous latent response variable. These thresholds are mapped on an unobserved continuous variable (Long, 1997), and, more importantly, they represent transitions from one category to another (commonly referred to as item difficulty). For example, consider an item that probes for citizen trust in government with three answer categories. Basically, the two thresholds between categories determine the difficulty of moving from one category to another. If we have similar respondents in two countries with the same position on the latent trait of trust, but different thresholds between item categories, then cross-national bias in response behaviour is present. Another well-known example here would be a psychological test on depression (cf. Janssen, 2011). Imagine one out of a battery of test items that asks respondents (men and women) how often they cried lately. Imagine also the test resulted in a higher depression score for women, mainly because many of them indicated that they cried often. However, crying has a lower threshold for women (e.g. Schaefer, 1988, cited in Janssen, 2011) resulting in a reduced difficulty to opt for a higher category – also known as differential item functioning. This means that if you would have a man and a woman with the same position on the latent trait of depression, women were more likely to cry when compared to men.

Within this framework, we define an item response model for each item: individual responses j (where j=1,…,J) for choosing category c are predicted using the cumulative probability ν_{ijkc} for each item i (where i=1,…,I) of a given respondent living in country k (where k=1,…,K). Thus it is a function of C –1 thresholds τ_{ic} (item difficulty) and the latent variable ξ_{jk} (that is the underlying latent trait we actually measure, for example 'trust in public institutions'), with the strength of the relationship between item and latent variable (the so-called discrimination parameter, or item loading) expressed in the models' coefficients λ_i (cf. Stegmueller, 2011). In other words, individuals' probability of choosing a higher item category is expressed as a result of their stronger 'trust' minus item difficulty. Hence formally, it can be expressed as follows:

$$\nu_{ijkc} = \tau_{ic} - \lambda_i \xi_{jk}. \tag{5}$$

For this graded response model, it has been found that it "[...] can be estimated with 250 respondents, but around 500 are recommended for accurate parameter estimates [when using a five point Likert scale]" (Reeve and Fayers, 2005: 70). However, here scholars need to be also aware of the respondents to parameter ratio; latent traits with many items require more respondents, than short scales. This conventional graded response model has been extended by Stegmueller (2011) to a multilevel mixture IRT model with item bias effects. Item bias (denoted as δ_{ik}) is expressed when item thresholds that are associated with the same score on the latent variable vary across countries, or when remembering our example above, across gender. It would mean that crossing a certain category for similar respondents is more difficult in country A than in country B. If this is the case, items are not equivalent across countries. Here, instead of testing and subsequently establishing (partial) equivalence (like one would do within a MGCFA framework), this approach corrects for measurement non-equivalence by explicitly modelling it. This is done by introducing discrete random effects for individual items to vary across mixtures m (where m=1,…,M) – these are groups, or more precisely latent classes, of countries that share unobserved heterogeneity in country item bias (denoted as η_{km}).[46] In such a model, item bias is allowed to vary across country mixtures that share unobserved heterogeneity in systematic responses behaviour. Or in other words, by introducing direct effects of these mixtures on items, we are able to explicitly model cross-national measurement non-equivalence.

Extending the graded response model, one has to make some changes in notation by first adding subscripts to equation (5), denoting the level of each parameter, with 1) items being nested in 2) individuals (where the latent concept 'trust' is located), nested in 3) countries (where the unobserved heterogeneity in country item bias is located). This yields a three-level model where we then also subtract the unobserved country item bias that varies across mixtures (cf. Stegmueller, 2011). Thus we get an unbiased cumulative response probability by specifying

$$\nu_{ijkc} = \tau_{ic} - \lambda_i^{(1)}\xi_{jk}^{(2)} - \sum_{m=1}^{M}\delta_{im}^{(1)}\eta_{km}^{(3)}. \tag{6}$$

When estimating this model, first the number of mixtures needs to be determined. This means that we need to figure out how many latent groups there are across the set of, for example, countries that share common characteristics in systematic country item bias. Hence the model from equation 6 should be estimated with an increasing number of mixtures. In a next step, scholars are able to compare fit

[46] Mixtures are composed of groups of countries that share the same posterior probability of responding (Vermunt and Magidson, 2005). These mixtures are specified to be categorical (using effect coding for model identification) yielding discrete random effects (Stegmueller, 2011).

measures (e.g. AICC, BIC; Log likelihood) of the different models to determine how many mixtures best fit their data.

In such a framework one can test for systematic country item bias by checking whether the estimates of item bias effects $\delta_i^{(1)}$ of single mixtures are significantly different from zero[47]. If this is the case, we would have strong evidence for the measurement non-equivalence of our items. Or in other words, this would mean that there exists systematic country item bias in response probability that stem from non-random threshold shifts across countries (cf. Stegmueller, 2011). Ignoring those differences would potentially yield biased estimates. Furthermore, this model specification allows us to add covariates to the model in equation (6) and subsequently estimate the 'true effects' of our independent variables of interest. This is possible because the multilevel mixture IRT approach explicitly models non-equivalence that stems from systematic country item bias, instead of using the common "two-step" approach of first testing for cross-national measurement equivalence, and then using (partial-)equivalent factors scores for subsequent estimations. The introduced IRT approach has the distinct advantage that it puts cross-national researchers in the position to explicitly correct for measurement equivalence, and estimate cross-national relationships within a single model.

6. Measurement Non-Equivalence in Practice

After having introduced both empirical techniques, in the following part of this study we will apply them to real life data. Our empirical examples come from cross-national public opinion surveys, as repeatedly used within comparative public administration research. The first example is on citizen satisfaction with public services, using continuous items, thus MGCFA is applied. The second example uses data on trust in public institutions, using ordinal items, and IRT modelling.

MGCFA and citizen satisfaction with public services

Recent years have seen an increasing interest in studying citizens' views and perceptions vis-à-vis public organisations. At the frontline of this development has been the examination of citizens' satisfaction with public services, including the interrelation with individual expectations (James, 2009; Morgeson, 2013; Van Ryzin, 200, 2013), its linkage with objective assessments of performance (Charbonneau and Van Ryzin, 2012; Shingler et al., 2008; Favero and Meier, 2013), or its propensity to facilitate citizens' trust in government (Vigoda-Gadot, 2007; Kampen, Van de

[47] For model identification, one has to set the item bias of one item to be zero – this is comparable to the MGCFA approach, where one item has to be utilized as the 'referent'.

Walle and Bouckaert, 2006). But also methodological considerations in measuring citizen satisfaction with public services have gathered pace (Herian and Tomkins, 2012; Van de Walle and Van Ryzin, 2011). Thus it can be seen that the study of citizen satisfaction with public services is of key interest for public administration scholars. A next desirable step would be the cross-national examination of theories of satisfaction in order to see whether they apply to different national contexts. Furthermore, linking individual data on citizen satisfaction with national, or regional, macro-level characteristics (such as the mode of delivery) would probe interesting findings regards micro-macro relationships. In pursuing such a research agenda, however, we first need to test whether citizen satisfaction, indeed, exhibits cross-national measurement equivalence.

Data

We use data from the European Consumer Satisfaction Survey (ECSS). Implemented on behalf of the European Commission, the ECSS was fielded in 2006. It covers all EU25 member countries[48] and a total of 11 different public services, and is thus one of the most comprehensive surveys on citizen satisfaction in Europe. Based on country stratifications according to region, urbanisation degree, gender, age and education, the ECSS makes use of a representative random sample for each service sector with a minimum of 500 respondents per sector and per country. For our example we use data from the electricity sector.

Here, service users have been asked to indicate their levels of satisfaction within this particular public service sector. More precisely they have been asked four questions that tap into their general levels of satisfaction with electricity services:

1. *Overall satisfaction (Sat Q1)*: "Overall, to what extent are you satisfied with [*supplier name*]? Please give me a score from 1 to 10 where 1) means that you are not satisfied at all, and 10) means that you are fully satisfied";
2. *Confirmation of expectations (Exp Q2)*: "If you compare what you expect from an electricity supplier and what you get from [*supplier name*], to what extent would you say that your requirements are met. Please give me a score from 1 to 10 where 1) means that your expectations are not met at all, and 10) means that your expectations are not only met but even exceeded";
3. *Satisfaction with service quality (Qual Q3)*: "I will read out a number of statements and would like you to give me, for each of them, a score where 1) means

[48] They include: Austria (AT), Belgium (BE), Cyprus (CY), Czech Republic (CZ), Denmark (DK), Estonia (EE), Germany (DE), Greece (GR), Finland (FI), France (FR), Hungary (HU), Ireland (IE), Italy (IT), Latvia (LV), Lithuania (LT), Luxembourg (LU), Malta (MT), Netherlands (NL), Poland (PL), Portugal (PT), Slovakia (SK), Slovenia (SI), Spain (ES), Sweden (SE), United Kingdom (UK).

that you totally disagree and 10) means that you totally agree: [*supplier name*] offers high quality services, overall";

4. *Price satisfaction (Price Q4)*: "I will read out a number of statements and would like you to give me, for each of them, a score where 1) means that you totally disagree, and 10) means that you totally agree: Overall, [*supplier name*]'s prices are fair, given the services provided".

Assessing cross-national measurement equivalence

For our study into citizens' satisfaction with electricity services, we first need to specify the model's factor structure (see Figure 7.4, here the lambdas for each item represent the highest and lowest lambda we find for all countries under analysis). All four items ought to tap the latent construct of citizens' satisfaction with electricity services. The first two items are quite similar, which is evident from their strong correlation ($r=0.803$; $p<0.000$). Thus we allow for a covariance between them. This can also be theoretically justified, since both items are directly probing for citizens' general satisfaction. Moreover, model assessments of individual countries without the covariance between them indicated that the model(s) would significantly and substantially improve by allowing a correlation between both items. This brings us to the measurement model as depicted in Figure 7.4. The figure also shows the factor loadings from the configural equivalent MGCFA model (highest and lowest country value). The model exhibits good measurement properties: all loadings are significantly different from zero, and load sufficiently strong on the latent trait of satisfaction.

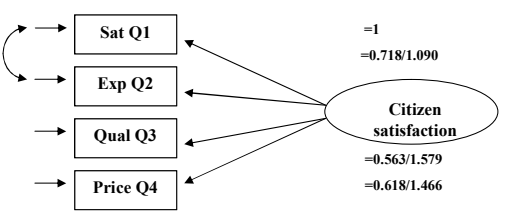

Figure 7.4: Measurement model of citizen satisfaction.

We test the measurement equivalence of citizens' satisfaction with their electricity services by using MGCFA. The measurement models were estimated using Mplus 6. We used a Maximum Likelihood Robust (MLR) estimator, which accounts for the non-normality of our items[49] (Muthén and Muthén, 2010: 533). Furthermore, we employed an estimation procedure that makes use of Full Information Maximum

[49] Q1: Skewness = −1.048, Kurtosis = 4.210; Q2: Skewness = −0.891, Kurtosis = 3.825; Q3: Skewness = −1.029, Kurtosis = 4.452; Q3: Skewness = −0.644, Kurtosis = 2.853.

Likelihood (FIML). FIML accounts for item non-response by taking all available data points into the estimation procedure, regardless whether there are missing cases, or not (see also Little and Rubin, 2002). In our case, item-non response was slightly above 5%.

For our analyses, we first determined the reference item to identify the scale of the latent variable. This choice has not been made arbitrarily, but is based on a procedure that sets the latent variable's variance to be 1 for all countries instead, and uses unstandardized modification index estimates to select the "most invariant item" (Sass, 2011), that is the item with the lowest overall modification index estimates – in our case, item Q1.[50] When it comes to the subsequent order of the test to assess our models' measurement equivalence, we employed a bottom-up strategy. This has been exemplified on Table 7.1 where the iterative process in equivalence testing is displayed. IT shows the respective model's fit to the data using the satora-bentler scaled chi-square, the model's degrees of freedom, and the RMSEA and CFI fit indices. More importantly, the change in chi square and standardized expected parameter change is displayed (STDYX EPC).

We start by assessing the configural equivalence of our measurement model, which means testing whether it has the same factorial structure within each country. We were able to establish the equivalence of our factor structure for all of the 25 countries under study. This means that within each country, all four items loaded significantly on a single factor. Moreover, fit indices of the multiple-group measurement model indicated that it fits the data well (see Table 7.1, model 0). Next we assessed the model's metric and scalar equivalence. The full metric model fits the data well, but it still can be improved substantially by releasing 3 constrained slopes (factor loadings). We were not able to establish full metric equivalence, since we found three countries with invariant factor loadings. However, by freeing the factor loadings for items Q3 and Q4, we can establish partial metric equivalence for all 25 countries. We can now meaningfully compare parameter estimates across all countries.

The next level of equivalence, full scalar, is much more difficult to satisfy. As depicted in Table 7.1, the full scalar model fits the data badly (model 6). However, it can be improved substantially by releasing 18 intercepts. After this, there were no further possibilities left for improving model fit. As we can see from Table 7.1,

[50] More specifically, Sass (2011: 354) proposes to set the variance of the unobserved latent variable to be equal to one for all groups. By this one would not need a referent, as the scale is already identified. On this basis, he recommends estimating a fully constrained measurement model and then using unstandardized overall modification indices for items' slopes and intercepts to select the referent.

Table 7.1: Equivalence tests for Citizens Satisfaction with Electricity Services; N = 13,155.

	Model specifications	χ^2	*df*	*RMSEA*	*CFI*	$\Delta\chi^2$	*STDYX EPC*
M0	**Configural equivalence**	29.60	25	0.035	0.999	–	–
M1	**Full metric equivalence**	233.14	97	0.067	0.981	–	–
M2	λ_{Q4}^{NL} released	219.87	96	0.064	0.983	13.27	−0.236
M3	λ_{Q4}^{PL} released	204.52	95	0.062	0.984	15.35	−0.176
M4	λ_{Q4}^{ES} released	189.35	94	0.059	0.986	15.17	−0.215
M5	λ_{Q3}^{LT} released	170.97	93	0.055	0.988	18.38	0.258
M6	**Full scalar equivalence**	1028.27	165	0.113	0.907	–	–
M7	τ_{Q3}^{SE} released	960.18	164	0.109	0.914	68.09	0.428
M8	τ_{Q2}^{CZ} released	904.00	163	0.106	0.920	56.18	−0.250
M9	τ_{Q1}^{LT} released	801.29	161	0.099	0.930	102.71	0.176
M10	τ_{Q2}^{AT} released	756.11	160	0.096	0.935	45.18	−0.238
M11	τ_{Q2}^{HU} released	711.41	159	0.093	0.939	44.70	−0.234
M12	τ_{Q2}^{DE} released	670.74	158	0.090	0.943	40.67	−0.228
M13	τ_{Q4}^{UK} released	631.84	157	0.087	0.947	38.91	0.260
M14	τ_{Q1}^{LT} released	598.72	156	0.085	0.951	33.12	0.191
M15	τ_{Q2}^{CY} released	564.50	155	0.082	0.954	34.21	0.155
M16	τ_{Q4}^{BE} released	536.48	154	0.080	0.957	28.02	0.237
M17	τ_{Q4}^{FI} released	509.24	153	0.078	0.959	27.25	0.248
M18	τ_{Q3}^{ES} released	481.46	152	0.075	0.962	27.78	−0.211
M19	τ_{Q4}^{MT} released	461.75	151	0.074	0.964	19.71	−0.219
M20	τ_{Q4}^{CZ} released	441.47	150	0.072	0.966	20.27	−0.213
M21	τ_{Q4}^{SK} released	423.68	149	0.070	0.968	17.79	−0.208
M22	τ_{Q4}^{SE} released	408.10	148	0.069	0.969	15.58	−0.215
M23	τ_{Q3}^{IE} released	391.89	147	0.067	0.971	16.21	−0.165
M24	τ_{Q3}^{NL} released	377.64	146	0.066	0.972	14.25	−0.155

Chi-square refers to the Satorra-Bentler scaled chi-square.

our final model displays an acceptable fit (model 24), with no fit index beyond what is generally considered to be an acceptable cut-off value. However, we are still not able to compare means across countries, since for meaningful comparisons we would need at least two items with the same invariant slopes *and* intercepts across countries (partial scalar equivalence). By freeing slopes *and* intercepts for items Q2 and Q4, we can now meaningful compare coefficients *and* latent country means for 19 countries. Yet, this excludes Ireland, Latvia, Lithuania, The Netherlands, Spain and Sweden, since they all have non-equivalent intercepts for items Q1 and Q3, which suggests that it is especially in those countries where items Q1 and Q3 function differently.

MGCFA: Does it matter?

In order to exemplify the biases comparative researchers may tap into when conducting cross-national analyses, we compare the results of our partial scalar equivalence model, with the status-quo in comparative research, simply computing a factor score for the measured concept form the pooled country data. We estimated simple country fixed effects linear regression models using 1) factors scores, and 2) the scores from our partial scalar equivalent MGCFA model. Figure 7.5 displays the results (using Austria, the country with the highest satisfaction scores, as reference). Differences between both approaches are striking. For example, using the standard factor score approach shows no significant difference between Austria and Luxembourg, while the results from our MGCFA model suggest that people in Luxembourg are significantly less likely to be satisfied with the services they receive; the same holds true for Slovenia and Ireland. The difference between both coefficients is the result of measurement non-equivalence. Furthermore, we can see that in many cases the MGCFA approach led to significantly different coefficients, resulting in a renewed country-order of levels of citizen satisfaction. Using pooled factor scores one may conclude that people living in the Czech republic, for example, are less satisfied with their electricity services than individuals from Slovakia, France, the United Kingdom, Poland, and Belgium. But when considering the MGCFA results it is pointed towards the opposite: Respondents in the Czech Republic are in fact *more* satisfied with their services than respondents form those other countries.

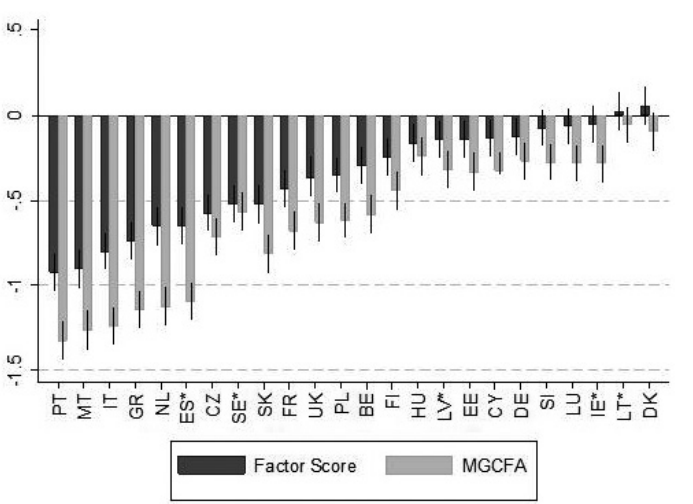

Figure 7.5: Country fixed effects and their respective 95% confidence intervals (citizen satisfaction).

** Denotes countries with invariant intercepts.*

These differences are the result of measurement non-equivalence, and not taking them into account can lead to biased results, and wrong theoretical conclusions.

IRT and Trust in Public Institutions

Trust in public institutions is regarded as an assessment of the performance, and procedural quality of these institutions. This trust is thought to influence citizens' willingness to obey or cooperate, and is as such an indicator of government's (political) legitimacy (Hooghe and Marien, 2011). Various scholars in public administration research have conducted empirical analysis into the determinants of institutional trust, and have looked at aspects such as performance, procedural quality, transparency, or frequency of contact (Grimmelikhuijsen and Meijer, 2014; Van de Walle, 2009; Van de Walle and Bouckaert, 2003; Van Ryzin, 2011). Trust in individual institutions is sometimes regarded as a reflection of not just specific institution's individual qualities, but also as a reflection of a wider propensity to trust public institutions (Mishler and Rose, 1997). Various cross-national analyses have been conducted in this regard, yet whether the concept of trust in public institutions travels across domestic borders was, to our knowledge, not subject to analyses. Studies of the longitudinal measurement-equivalence of trust in government in the US indicate that, indeed, the conception of trust changes over time, so as does people's individual response behaviour (Poznyak et al., 2013). In the following, we assess the cross-national measurement properties of citizen trust in public institutions, using the previously introduced IRT approach.

Data

For this part of our study, data from the World Value Survey (WVS) 2005 is used. WVS is a high-quality and well-known cross-national survey, established in 1981. It regularly surveys a representative sample of national populations across a very broad range of countries. It encompasses items on various theoretical concepts, including institutional trust. Using the WVS institutional trust inventory, Newton and Norris (2000) distinguish between trust in private and public institutions. The latter set of items is used for our IRT analysis, including trust in 1) the police, 2) the justice system, 3) the government, and 4) the civil service. More specifically, respondents were asked "*I am going to name a number of organisations. For each one, could you tell me how much confidence you have in them: is it a great deal of confidence, quite a lot of confidence, not very much confidence or none at all?*". This yields a set of four ordinal items that tap into individuals' trust in public institutions. Our analysis was conducted for the following 14 EU and/or OECD member countries: Bulgaria (BG), Canada (CA), East-Germany (DE-E), Finland (FI), Italy (IT), Norway (NO), Poland (PL), Romania (RO), Spain (ES), Slovenia (SI), Sweden (SE), Switzerland (CH), United States of America (US), and West Germany (DE-W).

Assessing cross-national measurement equivalence

To apply the introduced multilevel mixture IRT model on real-life data of people's trust in public institutions, we use the statistical software LatentGOLD version 4.5. In order to ease the estimation process, we draw a 50% random subsample for our analysis (see also Stegmueller, 2011 for a similar procedure). The hierarchical conceptualisation of our multilevel IRT model (items nested in individuals, nested in countries), enables us to account for item non-response in a transparent way. Assuming missingness at random (Little and Rubin, 2002), merely resulted in different cluster sizes at level-1. We found a share of 9% of missing data in our trust measure, and use a total of 8,317 respondents[51].

We determined the number of latent classes of countries that share common characteristics in individuals' response behaviour (mixtures) needed for our model, by estimating the IRT model as described in the previous part of this study (equation 6) with an increasing number of mixtures. In a next step we compared the fit measures of different models to determine which number of mixtures best fitted our data (see Table 7.2). We find that the model with a total of 10 mixture components yielded the best model fit[52], which can be illustrated by looking at the information theory-based fit measures, the Bayesian information coefficient (BIC), and the consistent Akaike's information coefficient (AICC).[53]

Table 7.2: Determining the number of mixture components for multilevel IRT mixture model.

Model	*No. of mixture components*	*Log-Likelihood*	*No. of parameters*	*AICC*	*BIC*
M1	3	−31,900	26	64,062	64,036
M2	4	−31,817	31	63,945	63,914
M3	5	−31,744	36	63,849	63,813
M4	6	−31,695	41	63,801	63,760
M5	7	−31,652	46	63,765	63,719
M6	8	−31,623	51	63,758	63,707
M7	9	−31,591	56	63,744	63,688
M8 (Final model)	10	−31,553	61	63,717	63,657
M9	11	−31,545	66	63,752	63,686

[51] After deleting those individuals that did not answer any of the trust items (1%).

[52] Estimating the same model using continuous random effects clearly provides no better fit to the data than using discrete random effects – results are available upon request.

[53] If we would select the final model merely on the basis of the log likelihood, we would select model 9 with a total of 11 mixtures. However, for our model we used the information theory-based fit measures, because they explicitly discriminate against increasing model complexity.

Table 7.3: Citizen trust in public institutions measurement model (model 8).

	Factor loading $\lambda_i^{(1)}$	*Standard Error*	*Threshold* τ_{i1}	*Threshold* τ_{i2}	*Threshold* τ_{i3}
Police	1.989*	0.054	2.777*	−1.434*	−4.601*
Justice system	2.637*	0.084	3.938*	−0.685*	−4.516*
Government	1.768*	0.049	4.543*	0.705*	−2.620*
Civil service	1.503*	0.043	4.370*	0.227*	−3.075*

** denotes p-value< 0.05.*

In a next step the properties of our measurement model are examined. Table 7.3 presents an overview of the actual factor loadings, and their accompanying item thresholds. First, we can see that our items exhibit good measurement properties: all items load significantly and strongly on one latent trait, which is trust in public institutions. Moreover, we can see that the thresholds clearly spread out across a wide range of our latent variable. Thus it can be concluded that our items load statistically and substantively significant on the latent trait, and that the thresholds cover a wide range of the latent variable, providing a precise measurement over a great share of the scale of trust in public institutions.

Now we turn to analysing the extent of systematic country item bias on individuals' response behaviour. Table 7.4 reports the coefficients and standard errors of item bias (δ_{ik}) for each survey item and mixture component. In order to reach model identification, item bias of the first item (trust in the police) was set to be zero (see also Stegmuellller, 2011). From the table we can clearly see that there exists severe country item bias. Also item bias is of the same direction for most countries (except for Bulgaria, and Finland and Norway). For all countries, item bias of at least one item is significantly different from zero. It highlights the crucial role systematic

Table 7.4: Item bias effects (model 8).

	Justice System		**Government**		**Civil Service**	
Country	*Coefficient*	*Standard Error*	*Coefficient*	*Standard Error*	*Coefficient*	*Standard Error*
SE	−0.824*	0.154	0.087	0.118	−0.754*	0.117
CH	−0.331*	0.141	−1.008*	0.110	−0.798*	0.104
DE-W	0.263	0.158	0.907*	0.124	0.867*	0.116
BG	0.438*	0.158	−0.525*	0.124	−0.494*	0.118
ES	−0.554*	0.141	−1.129*	0.113	−0.197	0.104
FIN, NO	−0.084	0.119	0.055	0.092	0.211*	0.086
IT, US	0.590*	0.109	0.493*	0.085	0.324*	0.079
DE-E	0.119	0.149	1.013*	0.117	0.642*	0.108
PL, RO, SI	−0.536*	0.104	−0.725*	0.082	−0.111	0.075
CA	0.919*	0.118	0.823*	0.092	0.311*	0.085

** denotes p-value< 0.05.*

country differences in response probability play for our measure of trust in public institutions. Item bias is the strongest in Switzerland, West-Germany and Canada. Looking at the effect directions of item biases, we have to bear in mind that survey items measured trust in a reverse manner – a low value indicated high levels of trust, while high values low levels. Thus we can see that respondents in Switzerland, for example, systematically overreport their trust in public institutions, while people living in the western part of Germany underreport their levels of trust. If researchers now simply compare responses from these countries without correction for country item bias, they will either systematically over- or underestimate peoples' trust in public institutions.

IRT: Does it matter?

To exemplify the systematic biases comparative scholars may encounter when analysing cross-national data, we used the results from our IRT model against the standard approach in the discipline, which is simply computing factor scores from pooled country data. Figure 7.6 reports the coefficients and accompanying 95% confidence intervals from linear regression models with country fixed effects. Norway – the country with the highest levels of trust – is used as reference category. From the table we can clearly see that simply ignoring country item bias in response probability can lead to misleading results. For example, when we look at the factor score coefficients for Switzerland, we may conclude that Switzerland is not significantly different from Norway. But when looking at the coefficients from the

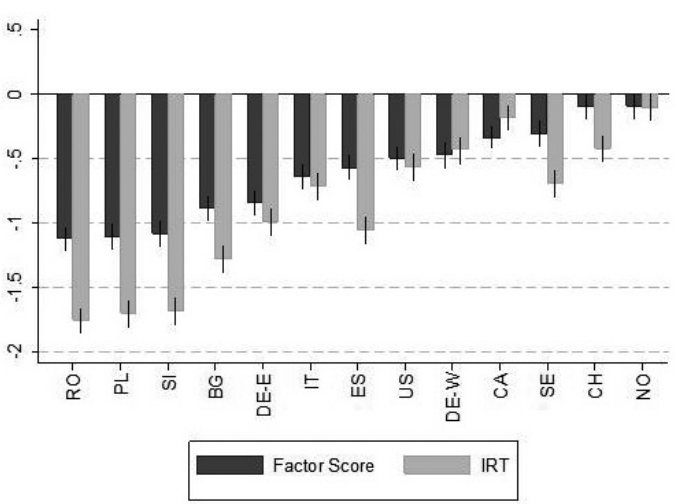

Figure 7.6: Country fixed effects and their respective 95% confidence intervals (trust in public institutions).

IRT approach used in this study, we see that individuals living in Switzerland are trusting their public institutions significantly less than people living in Norway. The difference between both coefficients is the result of systematic country item bias in individuals' item response probability. We, furthermore, find a different country-order of levels of trust. Using pooled factor scores one may conclude that people living in Spain, for example, are less trusting in public institutions than individuals from Italy, or East-Germany. The results from the IRT approach, however, suggests that mean levels of trust are actually *higher* in Italy compared to those countries. When looking at the position of Sweden, for example, a similar picture emerges. The pooled factor scores suggest the Swedes are more trusting in public institutions than respondents from Canada, West-Germany, and the US. However, after accounting for items bias effects through the illustrated IRT approach, a different picture comes into being. Now, respondents from Sweden are *less* trusting than people from those other countries. Again, these results come from systematic country item bias. Simply ignoring these differences can lead to invalid results, and wrong theoretical conclusions.

7. Conclusions: Measurement (Non-) Equivalence in Comparative Public Administration

Within comparative public administration survey research it is common practice to assume the equivalence of used latent traits and their accompanying survey items, possibly as a result of insufficient methodological training (cf. DeLorenzo, 2001; Gill and Meyer, 2000), or limited acquaintance with respective techniques. Researchers often simply pool items from different countries and subsequently utilize factor scores of the latent construct they measure. Seemingly, there is limited awareness among cross-national researchers within the field of public administration of the serious bias one may induce by pursuing such an estimation strategy (see Kim et al., 2012 for a notable exception). Our article has shown that for conducting meaningful cross-national analyses one needs to consider the cross-national equivalence of used survey measures. Estimating inferential models from comparative data without taking into account the possibility of measurement non-equivalence can lead to spurious results and misleading conclusions.

This article has presented two techniques to test and correct for measurement non-equivalence of comparative survey data in public administration research. Our empirical examples, indeed, show the biases one may get when pursuing the default approach of simply assuming the equivalence of measurement constructs. It was also exemplified that both concepts that we put under test (citizen satisfac-

tion, and trust in public institutions) did not exhibit cross-national measurement equivalence. Researchers who wish to conduct cross-national analyses using these concepts are best advised to account for their non-equivalence. But also comparative scholars who use measurement constructs whose cross-national measurement properties are unknown should do so, otherwise they risk biased results.

It becomes clear that for obtaining unbiased estimates, public administration scholars wishing to compare countries, or even regions (like the US states), are advised to test the equivalence of their measurement constructs. Otherwise they proceed in assuming equivalence, which can be, as we have shown, a very strong assumption. However, testing the non-equivalence assumption is straightforward and favourable, as it puts scholars in the position to test the geographical scope of their theories in a valid manner. This article has outlined two major techniques for doing so. We are aware of the increased difficulty in estimation this may bring along, however, the results from the last section have clearly shown that correcting for measurement non-equivalence is not a matter of fine-tuning estimates only of interest for methodologists, but of substantial importance when aiming to derive at approximately unbiased results that form the basis of our theoretical implications. Put simply, "*[…] doing serious comparative work entails additional effort*" (Pollitt, 2011: 124).

CHAPTER 8

Conclusions

1. Introduction to the Conclusions

In this dissertation we developed, and subsequently tested, theoretical predictions about citizens' deviations from the standard neoclassical model of rational decision-making within public infrastructure markets. In a first step, we have identified implicit, and often overlooked, microfoundations of competition and choice, and put forward two closely related assumptions that form their basis: (1) the Homo Oeconomicus assumption (*citizens act as rational customers within liberalised public service markets*), and (2) the homogenous rationality assumption (*all citizens in their role as customers act equally rational*). We have provided evidence – five separate empirical essays – which were summarised under the theoretical umbrella of these microfoundations. Each of the chapters provides stand-alone evidence, and thus can be regarded as an independent study. But the chapters also complement each other by providing a subsequent research flow, starting with investigating the Homo Oeconomicus assumption (chapter 3), to examining the homogeneous rationality assumption (chapter 4 and 5) and its consequences for citizens' welfare (chapter 6), as well as methodological considerations for the study of the microfoundations of competition and choice (chapter 7).

The major aim of this study was to increase our understanding of the behavioural microfoundations that form and have formed the basis of how public infrastructure services are provided. This is an important topic of theoretical and practical pedigree. We have provided evidence that questions the neoclassical model of citizens' market behaviour, and instead suggest a theoretical viewpoint of human behaviour in public infrastructure markets that acknowledges individuals' bounded rationality, and the differing degrees to which people – in particular those citizens who are potentially vulnerable as customers – are able to make choices within these markets. It was furthermore examined whether the identified bounded rationality in market behaviours among public service users is negatively related to the individual welfare they receive from these services. This has important implications for the equality criterion of the provision of services of general economic interest. However, we show that while the microfoundations of public infrastructure services seem not to work as predicted, markets can work after all. This means that once markets have a larger amount of citizens as customers that shop around for alternative providers, it creates market pressure for providers and thereby positive externalities for all groups of citizens. Thus, the distribution of some potential reform outcomes (such as lower prices) can spread more equally once effective choice is in place. This has important implications for the ways how academics and policy makers approach public infrastructure services' design and (soft) regulation by shifting the focus from the supply-side of public services to its demand-side.

2. Evidence on the Microfoundations of Competition and Choice

Have we found empirical support for the microfoundations of introducing competition and choice within liberalised public infrastructure services? We will answer our guiding research question by summarising the evidence that was obtained in this study. This will be done through an empirical assessment of the two microfoundations of competition and choice that were identified in the first chapter of this dissertation, the Homo Oeconomicus assumption and the homogeneous rationality assumption.

Homo Oeconomicus assumption

The Homo Oeconomicus assumption puts forward that citizens act as rational utility-maximising customers within liberalised public service markets. In other words, individuals will reflexively choose the service provider that provides them with the highest increase in their personal utility. We thus examined the Homo Oeconomicus assumption in the context of liberalised public infrastructure services by testing whether increasing choice among service suppliers leads citizens to switch providers after experiencing poor services (research question 1). In chapter 3, we use the theory of choice overload and apply it to citizens' reactions to failing public infrastructure services. Choice-overload states that increasing the number of alternatives reduces people's motivation to choose. We extend and test the theory of choice-overload by investigating whether or not increasing the number of providers of electricity services has detrimental effects on peoples' motivation to switch their provider after a service failure. We conducted a survey experiment were the number of service providers in a service failure scenario was randomly varied. Results show that increasing provider choice significantly reduces people's likelihood of switching away from a poor performing provider. These findings also hold when replicating the experiment with an independent online sample. Thus the results from chapter 3 indicate that increasing provider choice in public infrastructure markets causally influences people's motivation to not switch away from poor performing public services. Hence our first research question (*Do citizens in their role as customers become more likely to switch away from their current public service provider after experiencing a service failure and when choice is increased?)* can be answered with a no.

Homogenous rationality assumption

The homogenous rationality assumptions states that all citizens act equally rational in their market behaviours. It means that both, citizens who are potentially vulnerable as customers as well as their better-educated counterparts take decisions that are close to their individual optimum. Put differently, it is assumed that all citizens

in their role as customers are equally able to take up the opportunities of increased competition and choice by sending market signals to providers (i.e. complaining and switching). To investigate the homogenous rationality assumption we ask whether different layers of society – particular those who are better versus less well educated – differ in their abilities to send markets signals to service providers in public infrastructure markets were competition and choice have been implemented to varying degrees. Chapter 4 investigates the equality in citizens' choice behaviour (i.e. switching to another service provider) with regard to liberalised public services across twenty-five countries of the European Union. For our analysis we combined individual-level self-reported data about citizens' market behaviour with country-level information about the market characteristics of European public infrastructure services using hierarchical modelling techniques. Our findings suggest that potentially vulnerable and non-vulnerable groups of citizens do not send market signals in the very same manner under different degrees of choice, as assumed by the homogenous rationality assumption. Indeed, our analysis revealed that the gap between lesser and better educated public service users, in terms of actual switching behaviour, widens once a considerable degree of service liberalisation has been achieved. In other words, in service markets where there exist a larger number of service providers, less well educated people are less likely to have switched suppliers in the past two years. However, this choice-gap is apparent only once a certain threshold of choice is reached (more than eight providers) and was observed in the strongly liberalised mobile telephony sector, but not for less competitive services such as fixed telephony. Therefore, research question 2 (*Do different layers of society (particular those who are better versus less well educated) differ in their abilities to send markets signals to service providers in public service markets were choice has been implemented to varying degrees*?) can be partly answered with a yes.

In addition, chapter 5 has looked at whether citizens' complaint behaviour differs according to people's level of education (and age). While citizens who are potentially vulnerable as customers are indeed less likely to have complained about any aspect of their used service in the past two years, we do not find evidence that this gap is increasing over time. Unfortunately, the limited number of countries in our analysis (the EU 15 countries) did not permit us performing similar multilevel modelling techniques as in chapter 4, therefore we approached this research question in an alternative manner (looking at cross-sectional changes between 2000 and 2004). Results from this chapter add to our second research question by illustrating that it is not only that the less well educated citizens are less likely to switch among service providers, but they are also less likely to complain about the services they receive. Given previous works that have illustrated that especially this group of service users is least satisfied with the services they receive (for an

overview see Clifton et al., 2012), this is an important finding. It may point to the fact that potentially vulnerable citizens as customers are indeed locked-in with poor performing services. However, whether this relates to the process of service liberalisation remains to be tested.

Therefore, in chapter 6 we have investigated whether a lower frequency of market signalling that potentially vulnerable service users exhibit – as found in chapters 4 and 5 – relates to the welfare gains, or losses, they get when being potentially locked-in. While it is difficult to come up with an appropriate measure of individual welfare gains (a one Euro increase in welfare may have a different value across societal layers), we approach it from a subjective perspective by examining peoples' self-perceived affordability evaluations of the services they currently use. To address our research question, we performed a multilevel analysis of data combining individual affordability evaluations with country-level indicators about market structures of public infrastructure services across the EU-25 member states. In doing so, we related peoples' levels of self-perceived affordability with two measures of competition and choice: (1) the competitive market structure (whether choice and competition is available) – that is choice from the supply side – and (2) the competitive market functioning (whether choice is taken up by consumers) – that is choice from the demand-side. Our empirical findings suggest that less educated citizen-customers tend to experience lower levels of affordability of the services under analysis, when compared to those who are better educated. However, this gap is not significantly affected by markets' competitive structure. Instead, where this competitive structure is translated into a higher frequency of switching, the affordability gap between different socio-educational layers tends to be smaller and eventually disappears. Put simply, effective choice can reduce inequalities. Therefore we can answer our third research question (*Are there differences in individual welfare between better and less well educated citizens in public service markets were competition and choice have been implemented to varying degrees*?) with a yes. However, we have to note that the direction of the relationship is of the opposite direction as theoretically expected.

Cross-national measurement

Towards the end of this study we have examined how to study the microfoundations of competition and choice through cross-national survey data of citizen satisfaction with public infrastructure services, while simultaneously accounting for respondents' heterogeneous response behaviours across countries (*research question 4*). Within this methodological contribution, we have shown that for conducting meaningful cross-national analyses one needs to consider the cross-national equivalence of used survey measures. Estimating inferential models from

comparative data without taking into account the possibility of measurement non-equivalence can lead to spurious results and misleading conclusions. To illustrate this, we have estimated the biases one may get when pursuing the default approach of simply assuming the equivalence of measurement constructs by examining (1) citizen satisfaction with electricity services, and (2) respondent's trust in public institutions. We have done so by using two innovative measurement techniques that have previously received little or no attention in the public administration literature. Thus this chapter contributes to this dissertation by sketching out a methodological roadmap of how to study the microfoundations of competition and choice in future studies using cross-national public opinion data.

3. Discussion: Do Public Infrastructure Markets Work after All?

Our empirical results regarding the Homo Oeconomicus assumption provide evidence to suggest that choice-overload limits people's ability to respond to organisational failure because of the cognitive biases they face. Yet the ability of citizens to send market signals to poor performing service providers is one of the key assumptions put forward by theories of public service competition. In response, it is assumed that service providers would adjust their services to more closely match citizen's demands and preferences. But given the evidence of a choice-overload effect, the extent to which a match between citizen's preferences and demands and the offered services will be achieved can be questioned. In addition, the results of chapter 3 highlight that increasing provider choice in public service markets can potentially result in consumer inertia; this means that public service users could become locked-in to a suboptimal provider simply due to an overload of choices. Indeed, these findings stand in stark contrast with neo-classical economic thought of individuals acting as rational utility maximisers.

If we consider the theoretical rationale that in competitive markets providers adjust their goods and services in response to their customers' market signals, then our findings have severe implications for theories of public service delivery under competition. We have shown that potentially vulnerable service users in strongly marketised public service markets are less likely to switch, sending less market signals to providers then their better-off counterparts. Also they are less likely to complain about the services they receive. As a result, providers have fewer incentives to adjust the delivery of their services in accordance to their customers' needs and demands. This in turn may result in a decline in public service performance for those services used by potentially vulnerable citizens. In such a scenario, public infrastructure services may disproportionally benefit better-off service users. Indeed, this is what

Clifton et al. (2011a; 2011b) and Clifton, Díaz-Fuentes and Fernández-Gutiérrez (2014) have found in a series of studies (see also Fernández-Gutiérrez, 2011). Ultimately this may contributes to citizens potentially vulnerable as customers being locked-in with poor performing providers. This would represent a serious threat to the equality criterion of the European social model that underlies the establishment of services of general economic interest (see Clifton, Comin and Díaz-Fuentes, 2005; Héritier, 2001; 2002; Prosser, 2005).

However, the evidence we have obtained in this dissertation suggests that in countries where public infrastructure services' market competition is translated into a higher levels of national switching rates, the observed affordability gap between different socio-educational groups largely decreases. In other words, inequalities can be reduced and eventually disappear when more people switch within national markets. This suggests that public infrastructure providers more closely match their service offers – in terms of the prices – with citizen-customers' demands in markets with high national switching rates. We have applied a theory of marginal consumers to explain this relationship, by suggesting that once a critical mass of citizens in their role as customers switch among service providers, thereby creating market pressures from the demand-side, this creates positive externalities to all service users, including those who are potentially vulnerable in the marketplace[54]. Thus, it seems, that competitive markets work after all, but only once a competitive market structure is translated into higher switching rates.

4. A Future Research Agenda

A better knowledge of the demand-side perspective of public infrastructure services – and beyond – is required, if scholars and policy-makers aim to continue improving the functioning of these markets. Thus, this dissertation is not the end of a research process, but – hopefully – the end of the beginning. Therefore, inspired by the discussion and conclusion of our results, in this last section of our dissertation we will outline (1) limitations and future areas of our study, but also (2) practitioner recommendations for the delivery of public infrastructure services under

[54] Our findings also somewhat relate to Hirschman's (1970) ideas about the relationship between alert and inert consumers. He suggests that for competition and choice to work it is sufficient to have certain amounts of alert consumers in markets – or a right mix between inerts and alerts – that send market signals to providers. Those who are passive (the inert) will benefit from the trickle-down effects of the market signals of their more active counterparts. However, Hirschman's conception of the inert and alert is different to the concept of potential vulnerability. While inerts are primary seen as passive and somewhat apathic, the vulnerable have been defined as those "*[…] at a disadvantage in exchange relationships where that disadvantage is attributable to characteristics that are largely not controllable by them*" (Andreasen and Manning 1990, p.13).

competition and (3) provide, as well, an outlook into the study of a behavioural public administration.

Limitations and future research

Our empirical results open the door for interesting future research to follow-up on this dissertation. We particularly would like to highlight three areas of future research.

Firstly, our study has focussed on examining the microfoundations of competition and choice for public infrastructure services, mostly electricity and telecommunications services. As we outlined earlier, the choice of this empirical testing ground was of pragmatic nature: we needed an empirical sector which experienced liberalisation reforms but still has sufficient cross-national variation in competition and choice between national markets. Thereby, we were able to effectively tease out interesting relationships between competition, choice and individual-level behaviour. The focus on the public infrastructure sector, however, also limits the generalisability of our findings to other types of public services – though we can expect that cognitive mechanisms, such as choice overload work in an equally manner within other public service markets. Future studies are well advised to test our findings and theoretical propositions for public services such as education, health or social care. While these sectors have similarly experienced market-type reforms of introducing competition and choice into their supply, the degree to which citizens have real choice available varies across services. While overload of choices might be less likely to be a concern in these low choice environments, other cognitive biases in decision-making may do, such as the use of cognitive cues, or inertia more broadly.

Secondly, within this dissertation we have focused on examining reported behaviour (as well as self-perceived affordability evaluations) towards public infrastructure services. While we have little concern that people actually misreported their behaviours (reporting about whether one has send market signals is not necessarily subject to socially desirability bias, such as helping others, or donating blood is), still future works may look at citizens in their role as customers' real market behaviours, for example through examining people's actual spending behaviour (see for example Fernández-Gutiérrez, 2011), or even revealed choices in the context of randomised field experiments. This could not only bring-in individuals' revealed behaviour, but, more importantly, could rule-out any unobserved confounders that may have biased our results. Such a design could possibly be implemented in collaboration with local governments. A treatment group of municipalities may open-up competition between health centres, while a randomly selected

group of jurisdictions (the control group) keeps delivering their services though monopolistic state provision. In this way one can examine the counterfactual trend of marketisation reforms, thereby producing credible knowledge.

Thirdly, our findings may encourage further research on the boundary conditions of people's choice behaviour. Other factors than merely the number of service providers may also have bounded rational effects on people's choice behaviours. For example, the ways how different types of information about services is presented to service users may help them to overcome cognitive bias, however, it may also increases such biases. For example, prior research has shown that people are prone to more strongly react to poor, rather than to good, performance information – the so-called negativity bias. Here the way how information is framed (e.g. a mortality rate of 5% versus a survival rate of 95%) (see Kahneman and Tversky, 1979; also Olsen, 2014) might be crucial for how citizens react to poor services. But while there exist a growing literature on the effects of performance information on people's political voice behaviours (e.g. James, 2011; James and Moseley, 2014), we know relatively little about how they affect individuals' market behaviours, both switching but also complaining.

A related question would be whether people's market behaviours crowd-out their political voice. Within liberalised service markets citizens are not anymore treated as citizens in a civic-republican sense, but as customers of marketised services. While there indeed exist a literature that point to the emergence of a citizen-consumer (Clarke et al., 2007), we know virtually nothing about whether this change in people's role identities towards public services affect their political behaviours. For example, citizens are known to vote against incumbent politicians when public service performance goes below a certain threshold (Boyne et al., 2009; James and John, 2007). If people now have alternative accountability mechanisms available such as switching to better performing providers, will this decrease their political voice? Put simply, will we evidence an exit-voice trade-off (see also Hirschman, 1970; Dowding and John, 2012)? And moreover, will simply inducing citizens to a role identity of a customer rather than a citizen already affect their political voice behaviour, and thereby democracy? These are important questions that a future research agenda on the effects of market-oriented services from the citizens' perspective could tackle.

Practical relevance

Our findings are also of great importance for the way how public infrastructure services are delivered. We, by no means, argue to limit the amount of service providers within public infrastructure service markets. Chapter 6, for example, illustrates that

the value of competition and choice may come not from an increasing number of alternatives or simply competition *per se*, but from providing information, advice and help for citizen as customers' market behaviours. Thus, the evidence obtained in this study supports the increasing orientation of EU policies towards the demand-side of services, by focusing on empowering consumers in the market (e.g. European Commission, 2012), after having focussed on the supply-side during previous decades. In this regard, instruments for "soft regulation" and citizen empowerment could be used to make sure that there is an appropriate amount of switchers within national services that shop around for better services.

Moreover, we would argue that – from a normative point-of-view – potential vulnerable services users should receive greater attention in consumer protection policies in order to close the equality gap in their actual market behaviours. While markets may distribute resources in a seemingly equal manner once effective choice is in place, there are still disparities between societal layers in terms of their market behaviours. Therefore, empowering the vulnerable to actively participate in public infrastructure markets is of importance, especially when considering the equality criterion of the European social model that forms the basis for the creation of services of general economic interest. However, to date a great deal of policy attention has gone into reducing switching costs (e.g. number portability rates in the telephony sector) and other supply-side factors. Instead, demand-side factors such as search costs seem to be equally important, at least. Thus, for example, providing easily available information on service offers could be one possible response to an increased market complexity, and help to overcome cognitive biases. Another possible way to account for an increasing inequality in citizens' market behaviour could be the extension of service obligation contracts that protect the potentially vulnerable. Such contracts outline arrangements between non-public service providers and the regulating public bodies (for an overview see Cheung, 2005). They could, for example, include guidelines to strengthen the market situation of potentially vulnerable citizens in their role as customers across the EU.

A behavioural public administration

This study has examined a previously often overlooked unit of analysis in the study of public service reforms: citizens' evaluative judgements and behaviours vis-à-vis public services. Research into individuals' market behaviours requires an interdisciplinary perspective. Therefore, this study spans over the disciplines of public administration, psychology, and behavioural economics. It nicely fits into what Nobel laureate Daniel Kahneman (2013) has termed an applied behavioural science – making the processes within public administrations and public policy the unit of analysis of the behavioural sciences. In other words, we have looked into

individual behaviour and evaluative judgements by drawing upon recent advances in our understanding of the psychology and behaviour of individuals. That is the study of a behavioural public administration. Indeed, many public administration theories are macro or meso level conceptions. A behavioural public administration adds to their individual level underpinnings by approaching the behavioural and decision-making components of public administration from the demand-side. In other words, behavioural public administration re-examines the micro level assumptions of public administration theories. Thereby, a behavioural public administration stands in the tradition of Herbert Simon's earlier calls to more closely examine psychological theories of human decision-making and making them a central component in the study of public administration (Simon, 1955). However, these calls have stayed largely unanswered (Olsen, 2015). This dissertation contributes to the further development of such a research tradition by emphasising the micro level components of public administration theories that rest on often implicit and, moreover, seldom tested microfoundations.

5. Conclusion: Competition and Choice in the Delivery of Public Services

In past decades we have witnessed a massive restructuring of the way how public infrastructure services are delivered to citizens – the monopolist public provision of services to citizens was turned into competitive markets were multiple providers ought to compete for customers. Indeed, citizens in their new role as customers were expected to behave in the most rational manner by expressing their discontent via reflexively sending markets signals to poorly performing services by either complaining to, or switching between, public service providers. The suppliers of these services, in turn, were ought to adjust the value-for-money of their services accordingly, so that a long-run equilibrium could be reached between the needs and demands of citizens as customers and the services that have been provided to them. These market mechanisms rest on neo-classical economic thought of citizens acting as rational utility maximising calculators, and put forward assumptions about individual-level behaviour within public infrastructure markets. These are the microfoundations of competition and choice.

In our dissertation we have shown that some, but not all, citizens deviate from exhibiting such a rational market behaviour – they are bounded rational. Moreover, we show that there exist important differences between societal layers when it comes to sending market signals to providers. And despite being least satisfied, the potentially vulnerable are indeed less likely to compliant to, or switch between

providers of public infrastructure services. We have argued that this is so because of their limited capabilities in processing information and subsequent risk-assessments. We furthermore show that the switching-gap between these societal groups is larger in competitive markets where there exist greater opportunities to choose from. Or put simply, choice may increase inequalities in people's switching behaviour. However, our empirical analysis of whether these inequalities translate into welfare losses suggests that the right mix between switchers and non-switchers can potentially wipe-out inequalities between potentially vulnerable citizens as customers and their better-off counterparts. Thus, our dissertation puts the focus away from market regulation policies and puts the spotlight on enhancing customer empowerment and decreasing administrative burdens for exercising choice. Or put simply, for determining, and subsequently improving, the outcomes of liberalisation reforms, greater attention needs to be placed on the demand-side of public services. We hope this contribution has helped to shift a greater focus to the important role individual-level behaviours plays in determining aggregate level social outcomes of public services.

REFERENCES

Aberbach, J.D., & Christensen, T. (2005) Citizens and consumers. *Public Management Review, 7*(2), 225-246.

Acemoglu, D., & Verdier, T. (2000). The choice between market failures and corruption. *American Economic Review, 90*(1), 194-211.

Adang, E.M.M., & Borm, G.F. (2007). Is there an association between economic performance and public satisfaction in health care? *The European Journal of Health Economics, 8*(3), 279-285.

Adkins, N.R., & Ozanne, J.L. (2005). The low literate consumer. *Journal off Consumer Research, 32*(1), 93-105.

Aiken, L.S., & West, S.G. (1996). *Multiple regression: Testing and interpreting interactions*. London: Sage Publications.

Andreasen, A.R., & Manning, J. (1990). The dissatisfaction and complaining behaviour of vulnerable consumers. *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour, 3*, 12-20.

Appleby, J., Harrison, A., & Devlin, N.J. (2003). *What is the real cost of more patient choice?* London: Kings Fund.

Arrow, K. (1963). *Social Choice and Individual Values*. Second edition, New Haven: Yale University Press.

Arksey, H., & Glendinning, C. (2007). Choice in the context of informal care-giving. *Health & Social Care in the Community, 15*(2), 165-175.

Armstrong, M., & Sappington, D.E.M. (2006). Regulation, competition and liberalization, *Journal of Economic Literature, XLIV*, 325-366.

Ayeni, V. (2000). *Empowering the customer: the citizen in public sector reform*. London: Commonwealth Secretariat.

Bacchiocchi, E., Florio, M., & Gambaro, M. (2011). Telecom reforms in the EU: prices and consumers' satisfaction. *Telecommunications Policy, 35*(4), 382-396.

Bartels, K.P.R. (2013). Public encounters: the history and future of face-to-face contact between public professionals and citizens. *Public Administration, 91*(2), 469-483.

Bartlett, W., & Grand, J., le. (1993). *Quasi-markets and social policy*. Houndmills: Macmillan.

Bauby, P. (2008). L'européanisation des services publics. *Télescope, 14*(1), 11-22.

Baxter, K., Glendinning, C., & Greener, I. (2011). The implications of personal budgets for the home care market. *Public Money & Management, 31*(2), 91-98.

Behn, R.D. (1998). The new public management paradigm and the search for democratic accountability. *International Public Management Journal, 1*(2), 131-164.

Belloc, F., Nicita, A., & Parcu, P.L. (2013). Liberalizing telecommunications in Europe: Path dependency and institutional complementaries. *Journal of European Public Policy, 20*(1), 132-154.

Berkel, R., van, & Aa, P., van der. (2005). The marketisation of activation services: A modern panacea? Some lessons from the Dutch experience. *Journal of European Social Policy, 15*(4), 329-343.

Bismut, S. (2004). Telecommunications market outlook in new EU member states. *Communications & Strategies, 56*(4), 111-124.

Black, D. (1948). On the rationale of group decision-making. *The Journal of Political Economy, 56*(1), 23-34.

Bognetti G. & Oberman G. (2008) Liberalization and Privatization of Public Utilities: Origins of the Debate, Current Issues and Challenges for the Future. *Annals of Public and Cooperative Economics, 73*(3), 461-485.

Bollen, K.A. (1989). *Structural equations with latent variables*. New York: Willey.

Bollen, K.A. (2002). Latent variables in psychology and the social sciences. *Annual Review of Psychology, 53*, 605-34.

Botti, S., & Iyengar, S.S. (2004). The psychological pleasure and pain of choosing: When people prefer choosing at the cost of subsequent outcome satisfaction. *Journal of Personality and Social Psychology, 87*(3), 312-326.

Botti, S., & Iyengar, S.S. (2006). The dark side of choice: When choice impairs social welfare. *Journal of Public Policy and Marketing, 25*(1), 24-38.

Boyne, G., James, O., John, P. & Petrovsky, N. (2009). Democracy and Government Performance: Holding Incumbents Accountable in English Local Governments. *The Journal of Politics, 71*(04), 1273-1284.

Box, R.C., Marshall, G.S., Reed, B.J., & Reed, C.M. (2001). New public management and substantive democracy. *Public Administration Review, 61*(5), 608-619.

Brambor, T., Clark, W.R., & Golder, M. (2006). Understanding interaction models: Improving empirical analysis. *Political Analysis, 14*(1), 63-82.

Brau, R., Doronzo, R., Fiorio, C.V., & Florio, M. (2010). EU gas industry reforms and consumers' prices. *The Energy Journal, 31*(4), 163-178.

Brennan, C., & Coppack, M. (2008). Consumer empowerment: global context, UK strategies and vulnerable consumers. *International Journal of Consumer Studies, 32*(4), 306-313.

Brower, R.S., Mitchel Y. Abolafia, M.Y., & Carr, J.B. (2000). On improving qualitative methods in public administration research. *Administration & Society, 32*(4), 363-397.

Bruijn, H., de. (2007). *Managing performance in the public sector*. London: Routledge.

Bruijn, H., de, & Dicke, W. (2006). Strategies for safeguarding public values in liberalized utility sectors. *Public Administration, 84*(3), 717-35.

Buchanan, J.M., & Tullock, G. (1962). *The calculus of consent.* Ann Arbor: University of Michigan Press.

Buckley, J., & Schneider, M. (2003). Shopping for schools: How do marginal consumers gather information about schools? *Policy Studies Journal, 31*(2), 121-144.

Burden, R. (1998). *Vulnerable consumer groups: quantification and analysis*. London: Office of Fair Trading.

Burnham, T.A., Frels, J. K., & Mahajan, V. (2003). Consumer switching costs: A typology, antecedents, and consequences. *Journal of the Academy of Marketing Science, 31*(2), 109-126.

Butler, J.R. (1993). *Patients, policies and politics: Before and after working for patients*. Philadelphia: Open University Press.

Byrne, B.M., Shavelson, R.J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structure: The issue of partial measurement invariance. *Psychological Bulletin, 105*(3), 456-466.

Calero, J., Gil-Izquierdo, M., & Fernández-Gutiérrez, M. (2011). *Los costes del abandono escolar prematuro. Una aproximación a las pérdidas monetarias y no monetarias causadas por el abandono prematuro en España.* Madrid: Ministerio de Educación.

CEEP (European Centre of Employers and Enterprises Providing Public Services). (2010). *Public services in the European Union & in the 27 member states. Statistics, organization and regulations*. Available at: http://www.actionpublique.eu.

Chen, F.F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modelling, 14*(3), 464-504.

Chen, Y.-C., Shang, R.-A., & Kao, C.-Y. (2009). The effects of information overload on consumers' subjective state towards buying decisions in the internet shopping environment. *Electronic Commerce Research and Application, 8*(1), 48-58.

Cheung, A. (2005). Editorial on special issue: Service charters and customer empowerment in public services. *Public Management Review, 7*(3).

Christensen, T., & Lægreid, P. (2002). New public management: Puzzles of democracy and the influence of citizens. *Journal of Political Philosophy, 10*(3), 267-295.

Clarke, J., & Newman, J. (1997). *The managerial state: power, politics and ideology in the remaking of social welfare*. London: Sage.

Clarke, J., & Newman, J. (2007). What's in a name? New labour's citizen-consumer and the remaking of public services. *Cultural Studies, 21*(4-5), 738–757.

Clarke, J., Newman, J., Smith, N., Vidler, E., & Westmarland, L. (2007). *Creating citizen-consumers: changing publics and changing public services.* London: Sage Publications.

Clifton, J., & Díaz-Fuentes, D. (2010). Evaluating EU policies on public services: A citizens' perspective. *Annals of Public and Cooperative Economics, 81*(2), 281-311.

Clifton, J., & Díaz-Fuentes, D. (2008). The New Public Service Transnationals: Consequences for Labour. *Work Organisation, Labour and Globalisation, 2*(2), 178-194.

Clifton, J., Comín, F., & Díaz-Fuentes, D. (2005). Empowering Europe's citizens? Towards a charter for services of general interest. *Public Management Review, 7*(3), 417-443.

Clifton, J., Comín, F., & Díaz-Fuentes, D. (2006). Privatization in the European Union: Ideological, pragmatic, inevitable? *Journal of European Public Policy, 13*(5), 736-756.

Clifton, J., Comín, F., & Díaz-Fuentes, D. (eds.) (2007). *Transforming public enterprise in Europe and North America: Networks, integration and transnationalisation*. Houndmills: Palgrave-Macmillan.

Clifton, J., Díaz-Fuentes, D., & Fernández-Gutiérrez, M. (2014). The impact of socio-economic background on satisfaction: Evidence for policy-makers. *Journal of Regulatory Economics*, online first.

Clifton, J., Díaz-Fuentes, D., Fernández-Gutiérrez, M., James, O., Jilke, S., & Walle, S., van de. (2012). *Regulatory reform for services of general interest and trends in citizen satisfaction* (COCOPS research report number 3). Rotterdam: Erasmus University.

Clifton, J., Díaz-Fuentes, D., Fernández-Gutiérrez, M., & Revuelta, J. (2011a). Is market oriented reform producing a "two-track" Europe? Evidence for electricity and telecommunications. *Annals of Public and Cooperative Economics, 82*(4), 495-513.

Clifton, J., Díaz-Fuentes, D., Fernández-Gutiérrez, M., & Revuelta, J. (2011b). *The new regulation of public infrastructure services in the European Union. Challenges for territorial cohesion*. Paper presented at the European Regional Science Association Congress 2011, Barcelona.

Coe, C.K., & Brunet, J.R. (2006). Organizational report cards: Significant impact or much ado about nothing? *Public Administration Review, 66*(1), 90-100.

Conway, P., & Nicoletti, G. (2006). *Product market regulation in the non-manufacturing sectors of OECD countries: Measurement and highlights* (Economics Department Working Papers number 530). Paris: OECD.

Conway, P., Janod, V. & Nicoletti, G. (2005). Product Market Regulation in OECD Countries: 1998-2003. *OECD Economics Department Working Papers* 419, OECD Publishing.

Cookson, R., Dusheiko, M., Hardman, G., & Martin, S. (2010). Competition and inequality: Evidence from the English National Health Service 1991–2001. *Journal of Public Administration Research and Theory, 20*(1), 181-205.

Cooper, Z., G., Stephen, Jones, S., & McGuire, A. (2011). Does hospital competition save lives? Evidence from the English NHS patient choice reforms. *The Economic Journal, 121*(554), F228-F260.

Coulter, A., Maistre, N., le, & Henderson, L. (2005). *Patients' experience of choosing where to undergo surgical treatment*. Camden: Picker Institute.

Cseres, K.J. (2008). What has competition done for consumers in liberalised markets? *The Competition Law Review, 4*(2), 77-121.

Damman, O.C., & Rademakers, J. (2008). *Keuze-informatie in de zorg: Een internationale vergelijking van presentatiewijzen op internet*. Utrecht: Nivel.

Davidov, E., Meuleman, B., Billiet, J., & Schmidt, P. (2008). Values and support for immigration: A cross-country comparison. *European Sociological Review, 24*(5), 583-599.

Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., & Billiet, J. (2014). Measurement equivalence in cross-national research. *Annual Review of Sociology*, in press.

Davidov, E., Schmidt, P., & Billiet, P. (2011). *Cross-cultural analysis. Methods and applications*. New York, London: Routledge.

DeLorenzo, L. (2001). Stars aren't stupid, but our methodological training is: A commentary on Jeff Gill and Ken Meier's article "Public administration research and practice: A methodological manifesto". *Journal of Public Administration Research and Theory, 11*(1), 139-145.

Dixon, A., & Grand, J., le. (2006). Is greater patient choice consistent with equity? The case of the English NHS. *Journal of Health Services Research & Policy, 11*(3), 162-166.

Dixon, A., Grand, J., le, Henderson, J., Murray, R., & Poteliakhoff, E. (2007). Is the British National Health Service equitable? The evidence on socioeconomic differences in utilization. *Journal of Health Services Research & Policy, 12*(2), 104-109.

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *American Economic Review*, *100*(3), 1238-1260.

Dowding, K. (1992). Choice: its increase and its value. *British Journal of Political Science, 22*(3), 301-314.

Dowding, K., & John, P. (2007). The three exit, three voices and loyalty framework: A test with survey data on local services. *Political Studies, 56*(2), 288-311.

Dowding, K., & John, P. (2009). The value of choice in public policy. *Public Administration, 87*(2), 219-233.

Dowding, K., & John, P. (2011). Voice and choice in health care in England: Understanding citizen responses to dissatisfaction. *Public Administration, 89*(4), 1403-1418.

Dowding, K., & John, P. (2012). *Exits, voices, and social investment: Citizens' reactions to public services*. Oxford: Oxford University Press.

Dowding, K., John, P., Mergoupis, T., & Vugt, M., van. (2000). Exit, voice and loyalty: Analytic and empirical developments. *European Journal of Political Research, 37*(4), 469-495.

Downs, A. (1967). *Inside bureaucracy*. Boston: RAND.

Egeberg, M. (1995). Bureaucrats as public policy-makers and their self-interests. *Journal of Theoretical Politics, 7*(2), 157-167.

Enders, C.K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods, 12*(2), 121-138.

Epling, N.M. (2003). *Price discrimination amid heterogeneous switching costs: a competitive strategy of the long distance telephony fringe*. Available at SSRN: http://ssrn.com/abstract=2056043.

European Commission. (2003). *Green paper on services of general interest* (COM(2003) 270 final, 21 May). Available at: http://eur-lex.europa.eu/LexUriServ/site/en/com/2003/com2003_0270en01.pdf.

European Commission. (2004a). *EU competition policy and the consumer*. Luxembourg: European Commission.

European Commission. (2004). *White paper on services of general interest.* (COM(2004) 374 final, 12 May). Available at: http://eur-lex.europa.eu/LexUriServ/site/en/com/2004/com2004_0374en01.pdf.

European Commission. (2005). *Communication from the Commission to the Council and the European Parliament*: *Report on progress in creating the internal gas and electricity market* (COM(2005) 568, 15 November). Available at: http://ec.europa.eu/energy/gas_electricity/legislation/benchmarking_reports_en.htm.

European Commission. (2006). *Eurobarometer on services of general interest* (EB 65.3). Brussels: European Commission.

European Commission. (2008a). *How can behavioural economics improve policies affecting consumers?* DG Health and Consumer Protection, Brussels, 28th November.

European Commission. (2008b). *Questionnaire XIII implementation report. 13th Progress report towards a single European telecom market.* Brussels: European Commission.

European Commission. (2008c). *Report on progress in creating the internal gas and electricity market. Technical annex to the communication from the Commission to the Council and the European Parliament* (COM(2009) 115, 11 March 2009). Brussels: European Commission.

European Commission. (2009a). *14th Progress report on the single European electronic communications market 2008.* Brussels: European Commission.

European Commission. (2009b). *Flash Eurobarometer 243: Consumers' views on switching service providers.* Brussels: European Commission.

European Commission. (2010). *15th Progress report on the single European electronic communications market 2009*. Brussels: European Commission.

European Commission. (2012). *Commission staff working document on knowledge-enhancing aspects of consumer empowerment* (SWD(2012) 235, 19 July). Brussels: European Commission.

EUROSTAT. (2008). *European electricity market indicators 2006*. Available at: http://www.eds-destatis.de.

Falch, T., & S. Sandgren, S. (2006). *The effect of education on cognitive ability*. (Department of Economics Working Paper Series number 9). Trondheim: Norwegian University of Science and Technology.

Favero, N., & Meier, K.J. (2013). Evaluating urban public schools: Parents, teachers and state assessments. *Public Administration Review*, *73*(3), 401-412.

Fernández-Gutiérrez, M. (2011). *Aportaciones al análisis y la evaluación económica de los servicios públicos en la Unión Europea desde la perspectiva de los ciudadanos como consumidores*. PhD-thesis, Department of Economics, University of Santander.

Fernández-Gutiérrez, M., James, O. & Jilke, S. (2013). *Citizen Vulnerability and Satisfaction with Marketised Public Services*. Paper Presented at EGPA conference, 2013 in Edinburgh.

Ferrari, A.F., Pagani, L., & Fiorio, C.V. (2010). A two-step approach to analyze satisfaction data. *Social Indicators Research*, *104*(3), 545-554.

Fiorio, C.V., & Florio, M. (2009). *The reform of network industries, privatization and consumers' welfare: Evidence from the EU15* (DEAS Working Papers 2009-41). Milano: Università degli Studi di Milano.

Fiorio, C.V., & Florio, M. (2011). Would you say that the price you pay for electricity is fair? Consumers' satisfaction and utility reforms in the EU15. *Energy Economics, 33*(2), 178-187.

Fiorio, C.V., Florio, M., Salini, S., & Ferrari, P. (2007). European consumers' attitudes on services of general interest: accessibility, price and quality. In M. Marrelli, F. Padovano, & I. Rizzo (Eds.) *Servizi pubblici: Nuove tendenze nella regolamentazione, nella produzione e nel finanziamento*. Milano: Franco Angeli.

Florio, M. (2013). *Network industries and social welfare: The experiment that reshuffled European utilities*. Oxford: Oxford University Press.

FONDACA. (2008). *A citizens' evaluation of national regulatory authorities*. (Final report). Available at: http://www.rutlandcab.org.uk.

Fotaki, M., Roland, M., Boyd, A., McDonald, R., Scheaff, R., & Smith, L. (2008). What benefits will choice bring to patients? Literature review and assessment of implications. *Journal of Health Services Research & Policy, 13*(3), 178-184.

Friedman, M. (1955). The role of government in education. In R.A. Solo (Ed.), *Economics and the public interest* (pp. 123-144). New Brunswick, NJ: Rutgers University Press.

Gans, J.S. (2005). Protecting consumers by protecting competition: Does behavioural economics support this contention? *Competition and Consumer Law Journal, 13*(1), 40-50.

George, M., Graham, C., & Lennard, L. (2011). *Too many hurdles: information and advice barriers in the energy market*. Leicester: Eaga Charitable Trust, University of Leicester.

Gerpott, T., Rams, W., & Schindler, A. (2001). Customer retention, loyalty, and satisfaction in the German mobile cellular telecommunications market. *Telecommunications Policy, 25*(4), 249-269.

GESIS. (2008). *Weighting overview*. Available at: http://www.gesis.org/eurobarometer/survey-series/standard-special-eb/weighting-overview.

GESIS. (2010). *Sampling and fieldwork*. Available at: http://www.gesis.org/eurobarometer/survey-series/standard-special-eb/sampling-fieldwork.

GESIS. (2012). *Eurobarometer 65.3: Variable report*. Cologne: GESIS, Leibniz Institute for the Social Sciences.

Giddens, A. (2003). Introduction Neoprogressivism: A new agenda for social democracy. In A. Giddens (Ed.), *The progressive manifesto: New ideas for the centre-left* (pp. 1-34). Cambridge: Polity Press.

Giddens, A. (Ed.). (2003). *The progressive manifesto: New ideas for the centre-left*. London: Polity Press.

Gill, J., & Meier, K.J. (2000). Public administration research and practice: A methodological manifesto. *Journal of Public Administration Research and Theory, 10*(1), 157-199.

Giulietti, M., Waddams, C., & Waterson, M. (2005). Consumer choice and competition policy: A study of UK energy markets. *The Economic Journal, 115*(506), 949-968.

Glasby, J., & Littlechild, R. (2009). *Direct payments and personal budgets: Putting personalisation into practice*. London: Policy Press.

Glendinning, C. (2009). The consumer in social care. In I. Greener, R. Simmons, & M. Powell (Eds.), *The consumer in public services: Choice, values and difference*, (pp. 177-196). London: Policy press.

Godwin, R.K., & Kemerer, F.R. (2002). *School choice tradeoffs: Liberty, equity and diversity*. Austin: The University of Texas Press.

Gofen, A. (2012). Entrepreneurial exit response to dissatisfaction with public services. *Public Administration*, online first.

Gottfried, P.E. (2001). *After liberalism: Mass democracy in the managerial state*. Princeton: Princeton University Press.

Grand, J., le. (2005). Inequality, choice and public services. In A. Giddens, & P. Diamond (Eds.), *The new egalitarism*. London: Polity Press.

Grand, J., le. (2007). *The other invisible hand: delivering public services through choice and competition.* Princeton: Princeton University Press.

Greener, I., Simmons, R., & Powell, M. (2009). *The consumer in public services: choice, values and difference.* London: Policy Press.

Grimmelikhuijsen, S.G., & Meijer, A.J. (2014). Effects of transparency on the perceived trustworthiness of a government organization: Evidence from an online experiment. *Journal of Public Administration Research and Theory, 24*(1), 137-157.

Grosso, A., & Ryzin, G.G., van. (2012). Public management reform and citizen perceptions of the UK health system. *International Review of Administrative Sciences, 78*(3), 1-20.

Grzybowski, L. (2008). Estimating switching costs in mobile telephony in the UK. *Journal of Industry, Competition and Trade, 8*(2), 113-132.

Harvey, D. (2007). *A brief history of neoliberalism*. Oxford: Oxford University Press.

Hayek, F.A. (1944). *The road to serfdom.* London: Routledge.

Herian, M.N., & Tomkins, A.J. (2012). Citizen satisfaction survey data: A mode comparison of the derived importance-performance approach. *The American Review of Public Administration, 42*(1), 66-86.

Héritier, A. (2001). Market integration and social cohesion: the politics of public services in European regulation. *Journal of European Public Policy, 8*(5), 825-852.

Héritier, A. (2002). Public-interest services revisited. *Journal of European Public Policy, 9*(6), 995-1019.

Hirschman, A.O. (1970). *Exit, voice, and loyalty: Reponses to decline in firms, organizations, and states.* Cambridge: Cambridge University Press.

Hjorth, K., & Fosgerau, K. (2011). Loss aversion and individual characteristics. *Environmental and Resource Economics, 49*(4), 573-596.

Hogg, M.K., Howells, G., & Milman, D. (2007). Consumers in the Knowledge-Based Economy (KBE): What creates and/or constitutes consumer vulnerability in the KBE? *Journal of Consumer Policy, 30*(2), 151-158.

Holland, D., Lutz, C., Nonini, D.M., Barlett, L., Frederick-McGlathery, M., Gulbrandtsen, T.C., & Murillo Jr., E.G. (2007). *Local democracy under siege: Activism, public interests, and private politics.* New York: New York University.

Hood, C. (1991). A public management for all seasons. *Public Administration, 19*(1), 3-19.
Hood, C. (2006). Gaming in targetworld: The targets approach to managing British public services. *Public Administration Review, 66*(4), 515-521.
Horn, J.L., & McArdle, J.J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research, 18*(3), 117-144.
Hosmer, D.W., & Lemeshow, S. (2000). *Applied logistic regression*. New Jersey: Wiley-Interscience.
Howarth, D. (2008). Internal policies: The Commission defends the EU consumer. *Journal of Common Market Studies, 46*, 91-107.
Howarth, D., & Sadeh, T. (2010). The ever incomplete single market: Differentiation and the evolving frontier of integration. *Journal of European Public Policy, 17*(7), 922-935.
Howell, W.G. (204). Dynamic effects in means-tested, urban school voucher programs. *Journal of Policy Analysis and Management, 23*(2): 225-250.
Howell, W.G., Wolf, P.J., Campbell, D.E., & Peterson, P.E. (2002). School vouchers and academic performance: Results from three randomized field trials. *Journal of Policy Analysis and Management, 21*(2), 191-217.
Hox, J. (2002). *Multilevel analysis: Techniques and applications*. Mahwah, New Jersey, London: Lawrence Erlbaum.
Hox, J.J., de Leeuw, E.D., Brinkhuis, M.J.S., & Ooms, J. (2012). Multigroup and mutilevel apporaches to measurement equivalence. In S. Salzborn, E. Davidov, & J. Reinecke (Eds.), *Methods, theories, and empirical applications in the social sciences. Festschrift for Peter Schmidt* (pp. 91-96). Wiesbaden: Springer Verlag.
Hsieh, C.-T., & Urquiola, M. (2003). *When schools compete, how do they compete? An assessment of Chile's nationwide school voucher program* (NBER Working Paper number 10008). Cambridge: National Bureau of Economic Research.
Humphreys, P., & Simpson, S. (2008). Globalization, the 'competition state' and the rise of the 'regulatory' state in European telecommunications. *Journal of Common Market Studies, 46*(4), 849-874.
Hwang, M.I., & Lin, J.W. (1999). Information dimension, information overload and decision quality. *Journal of Information Science, 25*(3), 213-218.
Ilzkovitz, F., Dierx, A., & Sousa, N. (2008). *An analysis of the possible causes of product market malfunctioning in the EU: First results for manufacturing and service sectors* (Economic Papers 336). Brussels: European Commission.
Iyengar, S.S., & Lepper, M.R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology, 79*(6), 995-1006.
Jackson, D.L. (2003). Revisiting sample size and number of parameter estimates: Some support for the *N:* q hypothesis. *Structural Equation Modelling, 10*(1), 128-141.
James, O. (2009) Evaluating the expectations disconfirmation and expectations anchoring approaches to citizen satisfaction with local public services. *Journal of Public Administration Research and Theory, 19*(1), 107-123.
James, O. & Moseley, A. (2014). Does Performance Information about Public Services affect Citizen's Perceptions, Satisfaction and Voice Behaviour? Field Experiments with Absolute and Relative Performance Information. *Public Administration, 92*(2), 493-511.

James, O. (2011). Performance Measures and Democracy: Information Effects on Citizens in Field and Laboratory Experiments. *Journal of Public Administration Research and Theory, 21*(3):399-418.

James, O. (2003): *The Executive Agency Revolution in Whitehall: Public Interest versus Bureau-Shaping Perspectives*. Basingstoke: Palgrave MacMillan.

Janssen, R. (2011). Using a differential item functioning approach to investigate measurement invariance. In E. Davidov, P. Schmidt, & J. Billiet (Eds.), *Cross-cultural analysis: Methods and applications* (pp. 415-32). New York, London: Routledge.

Jilke, S., Meuleman, S. & Van de Walle, S. (2014). We Need to Compare, but How? Measurement Equivalence in Comparative Public Administration. *Public Administration Review,* online first.

Jilke, S. (2014). Choice and equality: Are vulnerable citizens worse-off after liberalization reforms? *Public Administration*, online first.

Jilke, S., Van Ryzin, G. & Van de Walle, S. (2014). Public Management Reform and Responses to Decline in Public Services: An Experimental Evaluation of Choice-Overload. *Paper presented at IRSPM conference 2014 in Ottawa.*

Jilke, S., & Walle, S., Van de. (2012). Two track public services? Citizens' voice behaviour towards liberalized services in the EU15. *Public Management Review, 15*(4), 465-476.

Jilke, S. 2009. Equal Access to Basic Services in African LDCs: The need for coherent, inclusive and effective Policy Frameworks. *zu|wuerfe discussion paper,* Friedrichshafen: Zeppelin University.

Johnson, T., Kulesa, P., Ik Cho, Y., & Shavitt, S. (2005). The relation between culture and response styles: Evidence from 19 countries. *Journal of Cross-Cultural Psychology, 36*(2), 264-277.

Jong, M.G., de, & Steenkamp, J.-B.E.M. (2010). Finite mixture multilevel multidimensional ordinal IRT models for large scale cross-cultural research. *Psychometrika, 75*(1), 3-32.

Jöreskog, K.G. (1971). Simultaneous factor analysis in several populations. *Psychometrika, 36*(4), 409-426.

Kahneman. D. (2013). *Foreword.* IN: Shafir, E. (ed.). The Behavioral Foundations of Public Policy. Princeton: Princeton University Press.

Kahneman, D. & Tversky, A. (1979): Prospect Theory: An Analysis of Decision Under Risk. *Econometrica, 47*(2), 263-292.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases.* Cambridge: Cambridge University Press.

Kampen, J.K., Walle, S., van de, & Bouckaert, G. (2006). Assessing the relation between satisfaction with public service delivery and trust in government: The impact of the predisposition of citizens toward government on evaluations of its performance. *Public Performance & Management Review, 29*(4), 387-404.

Kankaras, M., Vermunt, J.K., & Moors, G. (2011). Measurement equivalence of ordinal items: A comparison of factor analytic, item response theory, and latent class analysis approaches. *Sociological Methods & Research*, *40*(2), 279-310.

Kaul, I. & Conceicao, P. (Eds.) (2003). *Providing Global Public Goods: Managing Globalization*. Oxford: Oxford University Press.

Keller, K.L., & R. Staelin, R. (1987). Effects of quality and quantity of information on decision effectiveness. *Journal of Consumer Research, 14*, 200–213.

Kim, M.-K., Park, M.-C., & Jeong, D.-H. (2004). The effects of customer satisfaction and switching barrier on customer loyalty in Korean mobile telecommunication services. *Telecommunications Policy, 28*(2), 145-159.

Kim, S., Vandenaabeele, W., Wright, B.E., Bøgh Andersen, L., Cerase, F.P., Christensen, R.K., …, Vivo, P., de. (2012). Investigation the structure and meaning of public service motivation across populations: developing an international instrument and addressing issues of measurement invariance. *Journal of Public Administration Research and Theory*, *23*(1), 79-102.

King, G., Murray, C.J.L. Salomon, J.A., & Tandon, A. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review, 98*(1), 191-207.

Kline, R.B. (2011). *Principles and practice of structural equation modelling*. New York, London: The Guilford Press.

Kline, R.B. (2013). Exploratory and confirmatory factors analysis. In: Y. Petscher, & C. Schatsschneider (Eds.), *Applied quantitative analysis in the social sciences* (pp. 171-207). New York: Routledge

Knijn, T., & Verhagen, S. (2007). Contested professionalism payments for care and the quality of home care. *Administration & Society, 39*(4), 451-475.

Konisky, D.M., & Reenock, C. (2012). Case selection in public management research: Problems and solutions. *Journal of Public Administration Research and Theory*, doi: 10.1093/jopart/mus051.

Kremer, M. (2006). Consumers in charge of care: The Dutch personal budget and its impact on the market, professionals and the family. *European Societies, 8*(3), 385-401.

Kremer, M., & Tonkens, E. (2006). Authority, trust, knowledge and the public good in disarray. In J.-W. Duyvendak, T. Knijn, & M. Kremer (Eds.), *Policy, people and the new professional* (pp. 122-136). Amsterdam: Amsterdam University Press.

Lacireno-Paquet, N., Holyoke, T.T., Moser, M., & Henig, J.R. (2002). Creaming versus cropping: Charter school enrolment practices in response to market incentives. *Educational Evaluation and Policy Analysis, 24*(2), 145-158.

Ladd, H.F. (2002). School vouchers: A critical view. *The Journal of Economic Perspectives, 16*(4), 3-24.

Lako, C.J., & Rosenau, P. (2009). Demand-driven care and hospital choice. Dutch health policy toward demand-driven care: Results from a survey into hospital choice. *Health Care Analysis, 17*(1), 20-35.

Lambrecht, A., & Skiera, B. (2006). Paying too much and being happy about it: Existence, causes and consequences of tariff choice biases. *Journal of Marketing Research, XLIII*, 212-223.

Lane, J.-E. (2000). *New Public Management: An introduction*. London: Routledge.

Learmonth, M., & Harding, N. (2006). Evidence-based management: the very idea. *Public Administration, 84*(2), 245-266.

Lee, B.-K., & Lee, W.-N. (2004). The effect of information overload on consumer choice quality in an on-line environment. *Psychology and Marketing, 21*(3), 159-183.

Lee, G., Bennoit-Bryan, J., & Johnson, T.P. (2011). Survey research in public administration: Assessing mainstream journals with a total survey error framework. *Public Administration Review, 72*(1), 87-97.

Lee, J., Lee, J., & Feick, L. (2001). The impact of switching costs on the customer satisfaction-loyalty link: Mobile phone service in France. *Journal of Services Marketing, 15*(1), 35-46.

Lent, A., & Arend, N. (2004). *Making choices: How can choice improve local public services?* London: York Publishing Services.

Levi-Faur, D. (2004). On the 'net impact' of Europeanization: The EU's telecoms and electricity regimes between the global and the national. *Comparative Political Studies, 37*(1), 3-29.

Little, R.J.A., & Rubin, D.B. (2002). *Statistical analysis with missing data*. New Jersey: Wiley.

Lundsgaard, J. (2005). *Consumer direction and choice in long-term care for older persons, including payments for informal care: How can it help improve care outcomes, employment and fiscal sustainability?* (OECD Health Working Papers number 20). Paris: OECD.

Lyons, W.E., Lowery, D., & DeHoog, R.H. (1992). *The politics of dissatisfaction: Citizens, services, and urban institutions*. New York, London: M.E. Sharpe.

Marshall, M.N., Shekelle, P.G., Leatherman, S., & Brook, R.H. (2000). The public release of performance data. *JAMA: The Journal of the American Medical Association, 283*(14), 1866-1874.

Mayer, D.P., Peterson, P.E., Myers, D.E., Clarke Tuttle, C., & Howell, W.G. (2002). *School choice in New York City after three years: An evaluation of the School Choice Scholarships Program*. New York: Mathematica Policy Research.

McNutt, P. (2002). *The economics of public choice*. Aldershot: Edward Elgar.

Meier, K.J., & O'Toole, L.J. (2012). Subjective organizational performance and measurement error: Common source bias and spurious relationships. *Journal of Public Administration Research and Theory*, doi: 10.1093/jopart/mus057.

Meinow, B., Parker, M.G., & Thorslund, M. (2011). Consumers of eldercare in Sweden: The semblance of choice. *Social Science & Medicine, 73*(9), 1285-1289.

Meredith, W. (1993). Measurement invariance, factor analysis, and factoral invariance. *Psychometrika, 58*(4), 525-534.

Meredith, W. (1995). Two wrongs may not make a right. *Multivariate Behavioral Research, 30*(1), 89-94.

Meuleman, B. (2012). When are item intercept differences substantively relevant in measurement invariance testing? In S. Salzborn, E. Davidov, & J. Reinecke (Eds.), *Methods, theories, and empirical applications in the social sciences. Festschrift for Peter Schmidt* (pp. 97-104). Wiesbaden: Springer Verlag.

Meuleman, B., & Billiet, J. (2012). Measuring attitudes toward immigration in Europe: The cross-cultural validity of the ESS immigration scales. *Research and Methods, 21*(1), 5-29.

Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81-97.

Milsap, R.E., & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research, 39*(3), 479-515.

Mishler, W., & Richard, R. (1997). Trust, distrust and scepticism: Popular evaluations of civil and political institutions in post-communist societies. *The Journal of Politics, 59*(2), 418-451.

Moe, T.M. (1984): The New Economics of Organization. *American Journal of Political Science, 28*(4): 739-777.

Morgeson, F.V. III. (2013). Expectations, disconfirmation, and citizen satisfaction with US Federal Government: Testing and expanding the model. *Journal of Public Administration Research and Theory, 23*(2), 289-305.

Morris, J. (1993). *Independent lives? Community care and disabled people*. London: Macmillan.

Mullainathan, S., & Thaler, R. (2000). *Behavioural economics* (Working Paper Series Department of Economics 00-27). Cambridge: Massachusetts Institute of Technology.

Musset, P. (2012). *School choice and equity: Current policies in OECD countries and a literature review* (OECD Education Working Papers number 66). Paris: OECD.

Needham, C. (2003). *Citizen-consumers: New Labour's marketplace democracy*. London: Catalyst Forum.

Needham, C. (2007). *The reform of public services under New Labour: Narratives of consumerism*. New York: Palgrave Macmillan.

Newton, K., & Norris, P. (2000). Confidence in public institutions: Faith, culture, or performance? In S.J. Pharr, & R.D. Putnam (Eds.), *Disaffected democracies: What's troubling the trilateral countries?* (pp. 52-73). Princeton, NJ: Princeton University Press.

Niskanen, W.A. (1971). *Bureaucracy and public economics*. Aldershot: Edward Elgar.

Oberski, D.L. (2014). Evaluating sensitivity of parameters of interest to measurement invariance in latent variable models. *Political Analysis, 22*(1), 45-60.

OECD. (2008). *Enhancing competition in telecommunications: Protecting and empowering consumers*. Paris: OECD.

Oliver, R. L. (2010). *Satisfaction: A behavioral perspective on the consumer*. London, New York: M.E. Sharpe.

Olsen, J.P. (1988). Administrative reform and theories of organization. In B.G. Peters, & C. Campbel (Eds.), *Organizing governance: Governing organizations* (pp. 233-254). Pittsburgh: University of Pittsburgh Press.

Olsen, A. (2015). "Simon Said", We Didn't Jump. *Public Administration Review*, online first.

Olsen, A. (2014). Citizen (Dis)satisfaction: An Experimental Equivalence Framing Study. *Public Administration Review,* in press.

Osborne, D., & Gaebler, T. (1992). *Reinventing government: How the entrepreneurial spirit is transforming government*. New York: Penguin.

Ostrom, V., & Ostrom, E. (1971). Public choice: A different approach to the study of public administration. *Public Administration Review, 31*(2), 203-216.

O'Toole, L.J., & Meier, K.J. (2013). *Public management and performance: A theory of context*. Paper presented at the 11th Public Management Research Conference, Madison, Wisconsin, 20-22 June 2013.

Piurko, Y., Schwartz, S.H., & Davidov, E. (2011). Basic personal values and the meaning of left-right political orientations in 20 countries. *Political Psychology, 32*(4), 537-561.

Poggi, A., & Florio, M. (2009). Energy deprivation and dynamics and regulatory reforms in Europe: Evidence from household panel data. *Energy Policy, 38*(1), 253-264.

Poister, T., & Thomas, J.C. (2011). The effect of expectations and expectancy confirmation/disconfirmation on motorists' satisfaction with state highways. *Journal of Public Administration Research and Theory, 21*(4), 601-617.

Pollitt, C. (2011). Not odious but onerous: Comparative public administration. *Public Administration, 89*(1), 114-127.

Pollitt, C., & Bouckaert, G. (2004). *Public management reform: A comparative analysis*. Oxford: Oxford University Press.

Pollitt, C., & Bouckaert, G. (2011). *Public management reform: A comparative analysis – new public management, governance and the neo-Weberian state*. Oxford: Oxford University Press.

Portinga, Y.H. (1989). Equivalence of cross-cultural data: An overview of basic issues. *International Journal of Psychology, 24*(6), 737-756.

Poznyak, D., Meuleman, B., Abts, K., & Bishop, G.F. (2013). Trust in American government: Longitudinal measurement equivalence in the ANES, 1964-2008. *Social Indicators Research*, *118,* 741-758.

Prosser, T. (2005). *The limits of competition law – markets and public services*. Oxford: Oxford University Press.

Raju, N.S., Laffitte, L.J., & Byrne, B.M. (2002). Measurement equivalence: A comparison of methods based on confirmatory factor analysis and item response theory. *Journal of Applied Psychology, 87*(3), 517-529.

Ranganathan, C., Seo, D., & Babad, Y. (2006). Switching behavior of mobile users: Do users' relational investments and demographics matter? *European Journal of Information Systems, 15*, 269-276.

Reeve, B.B., & Fayers, P. (2005). Applying item response theory modelling for evaluating questionnaire item and scale properties. In: P. Fayers, & R.D. Hayes (Eds.), *Assessing quality of life in clinical trials: Methods of practice* (pp. 55-73). Oxford: Oxford University Press.

Reise, S.P., Widaman, K.F., & Pugh, R.H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin, 114*(3), 552-556.

Rensvold, R.B., & Cheung, G.W. (1998). Testing measurement models for factorial invariance: a systematic approach. *Educational and Psychological Measurement, 58*(6), 1017-1034.

Rouse, C. (1998). Private school vouchers and student achievement: An evaluation of the Milwaukee parental choice program. *The Quarterly Journal of Economics, 113*(2), 553-602.

Salucci, L., & Bickers, K. (2011). Exit, voice, and electoral turnover. *Urban Affairs Review, 47*(2), 155-182.

Samejima, F. (1969). *Estimation of latent ability using a response pattern of graded scores*. Richmond: William Byrd Press.

Samuelson, P.A. 1954. The Pure Theory of Public Expenditure. *The Review of Economics and Statistics, 36*(4): 387-389.

Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision-making. *Journal of Risk and Uncertainty, 1*(1), 7-59.

Saris, W.E., Satorra, A., & Sörbom, D. (1987). The detection and correction of specification errors in structural equation models. In C.C. Clogg (Ed.), *Sociological methodology* (pp. 105-129). San Francisco: Jossey-Bass.

Saris, W.E., Satorra, A., & Veld, W.M., van der. (2009). Testing structural equation models or detection of misspecifications? *Structural Equation Modelling, 16*(4), 561-582.

Sass, D.A. (2011). Testing measurement invariance and comparing latent factor means within a confirmatory factor framework. *Journal of Psychoeducational Assessment, 29*(4), 347-362.

Savas, E.S. (1977). An empirical study of competition in municipal service delivery. *Public Administration Review, 37*(6), 717-724.

Savas, E.S. (1987). *Privatization: The key to better government*. Chatham: Chatham House.
Savas, E.S. (2000). *Privatization and public-private partnerships*. New York: Chatham House.
Schaeffer, N.C. (1988). Application of item response theory to the measurement of depression. In C.C. Clogg (Ed.), *Sociological methodology* (pp. 271-308). San Francisco: Jossey-Bass.
Schmitt, C. (2013). The Janus face of Europeanisation: Explaining cross-sectoral differences in public utilities. *West European Politics, 36*(3), 547-563.
Schneider, M. Teske, P. & Marshall, M. 2002. *Choosing Schools: Consumer Choice and the Quality of American Schools*. New Haven: Princeton University Press.
Schneider, M., Teske, P., Marshall, M., & Roch, C. (1998). Shopping for schools: In the land of the blind, the one-eyed parent may be enough. *American Journal of Political Science, 42*(3), 769-793.
Schneider, M., Teske, P., Roch, C., & Marschall, M. (1997). Networks to nowhere: Segregation and stratification in networks of information about schools. *American Journal of Political Science, 41*(4), 1201-1223.
Schulten, T., Brandt, T., & Hermann, C. (2008). Liberalisation and privatisation of public services and strategic options for European trade unions. *Transfer: European Review of Labour and Research, 14*(2), 295-311.
Schwartz, B. (2005). *The paradox of choice: Why more is less*. New York: HarperCollins Publishers.
Shah, A.M., & Wolford, G. (2007). Buying behavior as a function of parametric variation of number of choices. *Psychological Science, 18*(5), 369-370.
Shingler, J., Loon, M.E., van, Alter, T.R., & Brider, J.C. (2008). The importance of subjective data for public agency performance evaluation. *Public Administration Review, 68*(6), 1101-1111.
Simon, H.A. (1972). Theories of Bounded Rationality. Chapter 8, In: McGuire, C.B. and Radner, R. (Eds.): *Decision and Organization: A Volume in Honor of Jacob Marschak*. Amsterdam: North-Holland Publication Co.
Simon, H.A. (1955). A Behavioral Model of Rational Choice. *Quarterly Journal of Economics, 69*, 99-188.
Six, P. (2003). Giving consumers of British public services more choice: What can be learned from recent history? *Journal of Social Policy, 32*(2), 239-270.
Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized latent variable modelling: Multilevel, longitudinal and structural equation models*. Boca Raton: Chapman & Hall.
Snow, J. (1855). *On the mode of Communication of Cholera*. London: John Churchill, New Burlington Street, England.
Sol, E., & Westerveld, M. (2007). The individual job seeker in the sphere of contractualism. *International Journal of Sociology and Social Policy, 27*(7/8), 301-310.
Steenkamp, J.-B.E.M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research, 25*(1), 78-107.
Stegmueller, D. (2011). Apples and Oranges? The problem of equivalence in comparative research. *Political Analysis, 19*(4), 471-487.
Stegmueller, D. (2013). How many countries for multilevel modelling? A comparison of frequentist and Bayesian approaches. *American Journal of Political Science, 57*(3), 748-761.
Stoker, G. & Moseley, A. (2010). *Motivation, Behaviour and the Microfoundations of Public Services*. London: 2020 Public Services Trust.

Teske, P., & Schneider, M. (2001). What Research can tell policymakers about school choice. *Journal of Policy Analysis and Management, 20*(4), 609-631.

Teske, P., Schneider, P., Mintrom, M., & Best, S. (1993). Establishing the micro foundations of a macro theory: Information, movers, and the competitive local market for public goods. *American Political Science Review, 87*(3), 702-713.

Thatcher, M. (1999). *The Europeanisation of regulation: The case of telecommunications* (Working Paper RSC 99/22). Florence: European University Institute.

Thomas, J.C. (1982). Citizen-initiated contacts with government agencies: A test of three theories. *American Journal of Political Science, 26*(3), 504-522.

Thomas, J.C., & Melkers, J. (1999). Explaining citizen-initiated contacts with municipal bureaucrats. *Urban Affairs Review, 34*(5), 667-690.

Thorpe, G.L., & Favia, A. (2012). *Data analysis using item response theory methodology: An introduction to selected programs and applications* (Psychology Faculty Scholarship Paper 20). Maine: University of Maine.

Tiebout, C.M. (1956). A pure theory of local expenditures. *The Journal of Political Economy, 64*(5), 416-424.

Timmermanns, D. (1993). The impact of task complexity on information use in multi-attribute decision making. *Journal of Behavioral Decision Making, 6*(2), 95-111.

Tullock, G. (1965). *The Politics of Bureaucracy*. Washington: Public Affairs Press.

Tummers, L., Jilke, S. & Van de Walle, S. (2013). Citizens in charge? An overview of the background, facilitators and pitfalls of introducing choice and competition in public services. IN: Divedi, Y.K., Shareef, M.A., Pandey, S.K. & Kumar, V. (eds), *Public Administration Reformation: Market Demand from Public Organizations*. New York: Routledge

Tummers, L., Bekkers, V., & Steijn, B. (2012). Policy alienation of public professionals: A comparative case study of insurance physicians and secondary school teachers. *International Journal of Public Administration, 35*(4), 259-271.

Tummers, L., Steijn, B., & Bekkers, V. (2012). Explaining willingness of public professionals to implement public policies: Content, context, and personality characteristics. *Public Administration, 90*(3), 716-736.

UK Department of Energy and Climate Change. (2013). *Helping customers switch: Collective switching and beyond*. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/253862/Helping_Customers_Switch_Collective_Switching_and_Beyond_final__2_.pdf.

Ungerson, C. (1997). Give them the money: Is cash a route to empowerment? *Social Policy & Administration, 31*(1), 45-53.

Ungerson, C. (2004). Whose empowerment and independence? A cross-national perspective on 'cash for care' schemes. *Ageing and Society, 24*(2), 189-212.

Vael, T., Vandekerckhove, S., Gyes, G., van, Roosbroek, S., van, Verhoest, K., & Coppin, L. (2008) *Liberalisation in services of general economic interest: A bottom up citizens' perspective. Analysis of the PIQUE survey*. Leuven: Higher Institute of Labour Studies.

Vandenberg, R.J., & Lance, C.E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, and recommendations for organizational research. *Organizational Research Methods. 3*(1), 4-69.

Van de Walle, S. & Boukaert, G. (2003). Public service performance and trust in government: The problem of causality. *International Journal of Public Administration, 29*(8/9), 891-913.

Van de Walle, S. & Roberts, A. (2008). Publishing performance information: An illusion of control? In W. van Dooren, & S. van de Walle (Eds.), *Performance information in the public sector: How is it used?* (pp. 211-226). Houndmills: Palgrave Macmillan.

Van de Walle, S. & Van Ryzin, G.G. (2011). The order of questions in a survey on citizen satisfaction with public services: Lessons from a split-ballot experiment. *Public Administration, 89*(4), 1436-1450.

Van de Walle, S. (2006). The impact of public service values on services of general interest reform debates. *Public Management Review, 8*(2), 183-205.

Van de Walle, S. (2008). What services are public? What aspects of performance are to be ranked? The case of "services of general interest". *International Public Management Journal, 11*(3), 256–274.

Van de Walle, S. (2009). Confidence in the criminal justice system: Does experience count? *British Journal of Criminology, 49*(3), 384-398.

Van de Walle, S. (2009). When is a service an essential public service? *Annals of Public and Cooperative Economics, 80*(4), 521-545.

Van de Walle, S. (2010). New Public Management: Restoring the public trust through creating distrust? In T. Christensen, & P. Lægreid (Eds.), T*he Ashgate research companion to New Public Management* (pp. 309-320). Alderson: Ashgate.

Van de Walle, S. & Hammerschmid, G (2011). The Impact oft he New Public Management: Challenges for Coordination and Cohesion in European Public Sectors. *Haldurskultuur – Administrative Culture, 12*(2), 190-209.

Van Ryzin, G.G. (2013). An experimental test of the expectancy-disconfirmation theory of citizen satisfaction. *Journal of Policy analysis and Management, 32*(3), 597-614.

Van Ryzin, G.G. (2011). Outcomes, process and trust of civil servants. *Journal of Public Administration Research and Theory, 21*(4), 745-760.

Van Ryzin, G.G. & Charbonneau, E. (2010). Public services use and perceived performance: An empirical note on the nature of the relationship. *Public Administration, 88*(2), 551-563.

Van Ryzin, G.G. (2006). Testing the expectancy disconfirmation model of citizen satisfaction with local government. *Journal of Public Administration Research and Theory, 16*(4), 599-611.

Vigoda-Gadot, E. (2007). Citizens' perceptions of politics and ethics in public administration: A five-year national study of their relationship to satisfaction with services, trust in governance, and voice orientation. *Journal of Public Administration Research and Theory, 17*(2), 285-305.

Vigoda-Gadot, E., Shoham, A., & Vashdi, D.R. (2010). Bridging bureaucracy and democracy in Europe: A comparative study of perceived managerial excellence, satisfaction with public services, and trust in governance. *European Union Politics, 11*(2), 289-308.

Vijver, F.J.R., van de, & Leung, K. (1997). *Methods and data analysis for cross-cultural research*. Newbury Park: Sage.

Vijver, F.J.R., van de. (1998). Towards a theory of bias and equivalence. *ZUMA Nachrichten Spezial*, *January 1998*, 41-66.

Vijver, F.J.R., van de. (2003). Bias and equivalence: Cross-cultural perspectives. In J.A. Harkness, F.J.R. van de Vijver, & P.Ph. Mohler (Eds.), *Cross-cultural survey methods* (pp. 143-156). New Jersey: Willey.

Walker, R.M., & Boyne, G.A. (2006). Public management reform and organizational performance: An empirical assessment of the UK Labour Government's public service improvement strategy. *Journal of Policy Analysis and Management, 25*(2), 371-393.

Wendt, C., Kohl, J., Mischke, M., & Pfeifer, M. (2010). How do Europeans perceive their healthcare system? Patterns of satisfaction and preference for state involvement in the field of healthcare. *European Sociological Review, 26*(2), 177-192.

West, P.A. (1998). Market-what market? A review of health authority purchasing in the NHS internal market. *Health Policy, 44*(2), 167-183.

Whittaker, T.A. (2012). Using the modification index and standardized expected parameter change for model modification. *Journal of Experimental Education, 80*(1), 26–44.

Wilson, C.M. (2012). Market frictions: A unified model of search and switching costs. *European Economic Review, 56*(6), 1070-1086.

Wilson, C.M., & Waddams-Price, C. (2010). Do consumers switch to the best supplier? *Oxford Economic Papers, 62*(4), 647-668.

Wilson, C.M., Garrod, L., & Munro, A. (2013). Default effects, transaction costs, and imperfect information. *Economic Letters, 119*(2), 213-215.

Wright, B.E., Manigault, L.J., & Black, T.R. (2004). Quantitative research measurement in public administration: An assessment of journal publications. *Administration & Society*, *35*(6), 747-764.

Wright, J.S.F. (2009). The regulatory state and the UK Labour Government's re-regulation of provision in the English National Health Service. *Regulation & Governance, 3*(4), 334-359.

Xavier, P. (2008). *Enhancing competition in telecommunications: protecting and empowering consumers.* Paris: OECD.

Zhu, L. (2012). Panel data analysis in public administration: Substantive and statistical considerations. *Journal of Public Administration Research and Theory*, doi: 10.1093/jopart/mus064.

APPENDIX A

Complete Results Chapter 4

Table 4.5: Individual-level descriptive statistics.

	Mobile services (N= 15,143)			Fixed services (N=13,422)		
	Mean	SD	Min; Max	Mean	SD	Min; Max
Education			1; 3			1; 3
Basic education	.151	.358	0; 1	.205	.205	0; 1
Secondary education	.488	.500	0; 1	.447	.447	0; 1
Higher education	.361	.480	0; 1	.348	.348	0; 1
Daily importance			1; 4			1; 4
Not at all important	.014	.118	0; 1	.013	.115	0; 1
Not very important	.109	.312	0; 1	.112	.315	0; 1
Fairly important	.322	.467	0; 1	.404	.491	0; 1
Very important	.555	.497	0; 1	.471	.499	0; 1
Easy to compare			1; 4			1; 4
Very difficult	.107	.309	0; 1	.105	.307	0; 1
Fairly difficult	.228	.420	0; 1	.253	.435	0; 1
Fairly easy	.403	.491	0; 1	.420	.494	0; 1
Very easy	.261	.439	0; 1	.222	.416	0; 1
Consumer interest protection			1; 4			1; 4
Very badly	.056	.230	0; 1	.055	.228	0; 1
Fairly badly	.222	.416	0; 1	.214	.410	0; 1
Fairly well	.608	.488	0; 1	.605	.489	0; 1
Very well	.114	.318	0; 1	.126	.332	0; 1
Age (Ref: 15-69 years old)	.050	.217	0; 1	.125	.330	0; 1
Gender (Ref: female)	.462	.499	0; 1	.448	.497	0; 1
Employment categories			1; 6			1; 6
Managers and professionals	.141	.348	0; 1	.139	.346	0; 1
Clerical workers	.220	.415	0; 1	.195	.397	0; 1
Self-employed	.065	.246	0; 1	.066	.248	0; 1
Working class	.148	.356	0; 1	.115	.319	0; 1
Unemployed	.059	.235	0; 1	.043	.203	0; 1
Not in Labour force	.367	.482	0; 1	.442	.497	0; 1
Place of residence			1; 3			1; 3
Rural village	.337	.473	0; 1	.359	.480	0; 1
Small/ middle town	.387	.487	0; 1	.375	.484	0; 1
Large town	.276	.447	0; 1	.266	.442	0; 1
Homeownership	.460	.498	0; 1	.513	.500	0; 1

Table 4.6: Country-level descriptive statistics.

	Mean	SD	Min, Max	Data source
Number of Service Providers (mobile)	7.080	2.929	3; 13	Own collection
Number of Service Providers (fixed)	3.440	2.399	1; 10	Own collection
Portability (in days) (mobile)	6.580	5.179	1; 20	European Commission
Portability (in days) (fixed)	9.080	6.855	0; 30	European Commission
Price (in Euros, PPP adjusted) (mobile)	.150	.058	.05; .27	European Commission
Price (in Euros, PPP adjusted) (fixed)	.366	.135	.19; .75	EUROSTAT
Concentration Ratio (mobile)	47.520	14.104	25; 94	EUROSTAT
Herfindahl Hirschman Index (fixed)	6186.360	2202.385	2717; 9791	EC
Subscribers (in thousands) (mobile)	19910.480	26042.187	347; 85700	EUROSTAT
Subscribers (in thousands) (fixed)	9229.720	13857.309	208; 54400	ITU

Table 4.7: Results.

	Mobile telephony		Fixed telephony	
	Model 1	Model 2	Model 1	Model 2
Control variables				
Daily importance (Ref: Not at all important)				
Not very important	1.397 (.272)	1.403 (.273)	1.333 (.331)	1.336 (.333)
Fairly important	1.661** (.314)	1.662** (.315)	1.234 (.298)	1.240 (.300)
Very important	2.066** (388)	2.071** (390)	1.471 (.354)	1.476 (.356)
Easy to compare (Ref: Very difficult)				
Fairly difficult	1.114 (.080)	1.116 (.081)	1.134 (.098)	1.132 (.098)
Fairly easy	1.263** (.090)	1.268** (.091)	1.117 (.095)	1.115 (.095)
Very easy	1.565** (.119)	1.570** (.119)	1.434** (.134)	1.433** (.134)
Consumer interest protection (Ref: Very badly)				
Fairly badly	.935 (.082)	.932 (.082)	.980 (.107)	.980 (.107)
Fairly well	.681** (.058)	.680** (.058)	.728** (.076)	.726** (.076)
Very well	.731** (.074)	.730** (.074)	.684** (.084)	.682** (.084)
Age (Ref: 15-69 years old)	.463** (.054)	.465** (.054)	.747** (.068)	.744** (.068)
Gender (Ref: female)	1.091* (.044)	1.089* (.044)	.908 (.045)	.908 (.045)
Employment category (Ref: Managers and workers)				
Clerical workers	1.049 (.070)	1.046 (.070)	.993 (.081)	.996 (.081)
Self-employed	1.117 (.103)	1.117 (.103)	1.115 (.124)	1.120 (.124)
Working class	.987 (.076)	.983 (.075)	1.174 (.112)	1.182 (.113)
Unemployed	1.091 (.106)	1.092 (.106)	.939 (.125)	.944 (.126)
Not in Labour force	1.007 (.064)	1.007 (.064)	.864 (.068)	.867 (.068)
Place of residence (Ref: large town)				
Small/ middle town	.951 (.047)	.949 (.047)	.826** (.051)	.825** (.051)
Rural village	.917 (.048)	.914 (.048)	.835** (.053)	.836** (.053)
Homeownership	.772** (.035)	.772** (.035)	.912 (.051)	.912 (.051)
Number portability (in days)	.976 (.018)	.976 (.018)	1.018 (.011)	1.011 (.012)
Subscribers (in thousands)	1.000 (.000)	1.000 (.000)	1.000 (.000)	1.000 (.000)
Price (in Euros, PPP adjusted)	.987 (.017)	.987 (.017)	1.010 (.006)	1.011 (.006)
Market concentration	.982* (.008)	.982* (.008)	.999** (.000)	.999** (.000)

APPENDIX B

Complete Results Chapter 6

Table 6.8: Continued results fixed telephony.

FIXED TELEPHONY				
	Model 1 Market competitive structure	Model 2 Market comp. structure + interaction	Model 3 Market competitive functioning	Model 4 Market comp. functioning + interaction
Service user	1.917** (.106)	1.915** (.106)	1.917** (.106)	1.918** (.106)
Access to services	3.113** (.244)	3.112** (.244)	3.113** (.244)	3.113** (.244)
Daily importance (Ref: Very important)				
Fairly important	.790** (.032)	.789** (.032)	.790** (.032)	.788** (.032)
No very important	.525** (.029)	.524** (.029)	.525** (.029)	.523** (.028)
Not at all important	.345** (.025)	.344** (.025)	.345** (.025)	.343** (.025)
DK	.382** (.084)	.382** (.084)	.382** (.084)	.384** (.085)
Political orientation (Ref: left leaning)				
Centre	1.144** (.053)	1.139** (.053)	1.144** (.053)	1.139** (.053)
Right leaning	1.101 (.056)	1.110 (.056)	1.101 (.056)	1.100 (.056)
DK	.854** (.043)	.852** (.043)	.854** (.043)	.852** (.043)
Age	.950** (.005)	.950** (.005)	.950** (.005)	.949** (.005)
Age squared	1.000** (.000)	1.000** (.000)	1.000** (.000)	1.000** (.000)
Place of residence (Ref: rural village)				
Small/ middle town	1.107* (.044)	1.106* (.044)	1.107* (.044)	1.107* (.044)
Large town	1.160** (.053)	1.159** (.053)	1.160** (.053)	1.160** (.053)
Native born	.917 (.068)	.919 (.068)	.917 (.068)	.922 (.069)
Social class (Ref: Employers)				
High-level nonmanual	1.025 (.119)	1.022 (.119)	1.025 (.119)	1.020 (.119)
Medium-level nonmanual	1.029 (.104)	1.032 (.105)	1.029 (.104)	1.030 (.105)
Low-level nonmanual	.914 (.074)	.912 (.074)	.914 (.074)	.913 (.074)
Self-employed in primary sector	1.159 (.212)	1.157 (.212)	1.159 (.212)	1.147 (.210)
Workers	.785** (.067)	.785** (.067)	.785** (.067)	.782** (.066)
Unemployed	.680** (.069)	.679** (.069)	.680** (.069)	.677** (.068)
Not working	.813** (.065)	.812** (.064)	.813** (.065)	.809** (.064)
Gender (Ref: female)	1.116** (.039)	1.115** (.039)	1.116** (.039)	1.116** (.039)
Household size (Ref: 1 person)				
2 Persons	1.133** (.054)	1.132** (.054)	1.133** (.054)	1.128* (.054)
3 Persons	1.142* (.066)	1.140* (.066)	1.141* (.066)	1.136* (.066)
4 Persons or more	1.137* (.064)	1.134* (.063)	1.137* (.063)	1.130* (.063)
Homeownership	1.138** (.049)	1.139** (.050)	1.138** (.049)	1.140** (.050)
Market size	1.000 (.000)	1.000 (.000)	1.000 (.000)	1.000 (.000)
GINI	.969 (.026)	.969 (.025)	.969 (.026)	.969 (.026)
Years since liberalisation	1.003 (.035)	1.003 (.035)	1.010 (.029)	1.010 (.029)
Price dynamics	1.025 (.033)	1.025 (.033)	1.026 (.034)	1.026 (.034)

Table 6.9: Continued results electricity.

ELECTRICITY				
	Model 1 Market competitive structure	Model 2 Market comp. structure + interaction	Model 3 Market competitive functioning	Model 4 Market comp. functioning + interaction
Service user	1.202 (.133)	1.203 (.133)	1.179 (.130)	1.181 (.130)
Access to services	2.439** (.196)	2.442** (.197)	2.467** (.197)	2.467** (.197)
Political orientation (Ref: left leaning)				
Centre	1.077 (.046)	1.076 (.046)	1.090* (.045)	1.090* (.045)
Right leaning	1.054 (.050)	1.054 (.050)	1.068 (.049)	1.068 (.049)
DK	.842** (.040)	.842** (.040)	.851** (.040)	.852** (.040)
Age	.953** (.005)	.953** (.005)	.951** (.005)	.952** (.005)
Age squared	1.000** (.000)	1.000** (.000)	1.000** (.000)	1.001** (.000)
Place of residence (Ref: rural village)				
Small/ middle town	1.063 (.039)	1.062 (.039)	1.070 (.038)	1.068 (.038)
Large town	1.122** (.047)	1.121** (.047)	1.144** (.048)	1.142** (.047)
Native	.856* (.056)	.856* (.056)	.845* (.055)	.848* (.055)
Social class (Ref: Employers)				
High-level nonmanual	1.166 (.122)	1.164 (.122)	1.199 (.124)	1.202 (.124)
Medium-level nonmanual	1.233* (.112)	1.234* (.112)	1.292* (.115)	1.300** (.116)
Low-level nonmanual	1.097 (.080)	1.097 (.080)	1.123 (.080)	1.123 (.080)
Self-employed in primary sector	.917 (.148)	.917 (.147)	.967 (.153)	.970 (.154)
Workers	.868 (.066)	.868 (.066)	.904 (.068)	.906 (.068)
Unemployed	.781** (.070)	.781** (.070)	.798* (.071)	.800* (.071)
Not working	.953 (.067)	.952 (.067)	.980 (.068)	.981 (.068)
Gender (Ref: female)	1.178** (.038)	1.177** (.040)	1.182** (.038)	1.179** (.038)
Household size (Ref: 1 person)				
2 Persons	1.018 (.038)	1.017 (.045)	.995 (.043)	.994 (.043)
3 Persons	.924 (.049)	.924 (.049)	.897* (.047)	.900* (.047)
4 Persons or more	.885* (.045)	.884* (.045)	.861** (.043)	.860** (.043)
Homeownership	1.216** (.048)	1.217** (.048)	1.219** (.047)	1.220** (.047)
Market size	.000** (.000)	.000** (.000)	.000** (.001)	.000** (.001)
GINI	.953 (.036)	.954 (.036)	.961 (.031)	.961 (.031)
Years since liberalisation	1.109 (.082)	1.108 (.082)	.921 (.085)	.920 (.085)
Price dynamics	.933** (.022)	.933** (.022)	.925** (.021)	.925** (.021)

SUMMARY

Research Questions of the Study

In past decades, public infrastructure markets have been opened up for competition among different service providers (Conway, Janod and Nicoletti, 2005), creating a situation in which public sector organisations have to compete with other suppliers in order to stay in business, and thus can no longer rest on their monopolist-status. This has not been without a reason. Inspired by creation of the European single market, as well as by public choice theory, it was assumed that increasing competition between service providers would provide market incentives to produce better services for lower prices in order to sustain on the market. In other words, after successful liberalisation reforms, service providers would need to compete for customers. This logic implies that citizens in practice would have the choice to exit service providers (Hirschman, 1970, pp. 21-25). Indeed, a key attribute in the provision of public infrastructure services is that the classical exit option of completely withdrawing from the service in question is often not feasible, too difficult or associated with extremely high costs (see Clifton et al., 2012). For instance, service users who choose to exit electricity or gas services would likely face severe effects on their individual wellbeing – especially in the winter. Here, choice, that is switching between different (public or private) service providers as a sub-form of exit (Dowding and John, 2012), becomes important. Thus through switching, or implicitly threatening to switch via complaints (what Hirschman [1970] calls "voice"), citizens send market signals to providers which are expected to create incentives to deliver greater value for money in order to keep existing customers, as well as attract new ones. As a results, it is assumed that a long-run equilibrium would be achieved between citizens' demands and preferences, and the price and quality of the offered services.

This neo-classical economic perspective on public service delivery under competition tends to assume that inducing competition into public services and increasing the number of service providers to choose from would result in an optimal allocation of available resources. Or in other words, public services would become cheaper and better. This rests on an important set of micro-level assumptions, the microfoundations of competition and choice.

The theoretical accounts on the introduction of competition and choice into public service delivery make important assumptions about individual-level motivations and behaviours, in particular stemming from the idea that individuals are rational, utility-maximizing calculators (see also Stoker and Moseley, 2010). Indeed, macro-level reforms of introducing market elements into the delivery of public services are built around the conception of the *Homo Oeconomicus*. However, there exists

limited awareness, and thus even less concrete empirical evidence whether these microfoundations of competition and choice get it right. For this study we have identified two important micro-level assumptions behind the introduction of competition and choice into public service delivery: The first on is that **citizens act as rational customers within liberalised public service markets** (*Homo Oeconomicus assumption*). The extension of citizen choice assumes that citizens identify and subsequently choose the service offer that best matches their needs and demands (European Commission, 2004; Stone, 2005). Or in other words, they reflexively choose the one service provider that provides them with the highest increase in their personal utility. Second, this also implies that **all citizens in their role as customers act equally rational.** This is what we call the *homogenous rationality assumption*. It would mean that all citizens in their role as customers will take those decisions that are close to their individual optimum, and hence increase their welfare homogenously. Put differently, it is assumed that all citizens in their role as customers are equally able to take up the opportunities of increased competition and choice. Their behaviour as customers is equally rational for all groups of citizens. Thus, liberalisation reforms have been implemented by not only assuming that all groups of citizens act equally rational, but also that increases in individual welfare would be allocated evenly among them (European Commission, 2004; see also Clifton et al., 2011a).

Within this dissertation we have examined the limitations of rational accounts about how citizens behave within public service markets. By drawing upon behavioural research on individual decision making we offer an alternative view of how citizens actually behave within liberalised public service markets and test it empirically. Concretely, we have investigated whether citizens indeed make rational decisions in public service markets where competition and choice have been inserted. In doing so, we have tested a set of research questions that are of great theoretical and practical pedigree for the way of how contemporary public service delivery is organised. We first examined the Homo Oeconomicus assumption by asking whether increasing choice leads citizens to switch providers after experiencing poor services (**chapter 3**). Next we addressed the homogenous rationality assumption by looking at whether different layers of society (particular those who are better versus less well educated) differ in their abilities to switch between service providers in public service markets were choice has been implemented to varying degrees, and whether we see similar patterns for their complaint behaviour (**chapters 4 and 5**). In the following, we have looked at whether different ways of how citizens in their role as customers act within liberalised public service markets may affect the degree to which they "benefit" from these market arrangements, and whether this also varies among different layers of society. Thus we have investigated whether dispari-

ties between better and less well educated service users are more, or less, strong in markets where choice and competition have been introduced to varying degrees (**chapter 6**). At the end of the study we have provided a methodological outlook of how to study the microfoundations of competition and choice through survey data of citizen satisfaction with public infrastructure services, while simultaneously accounting for respondents' heterogeneous response behaviours across countries (**chapter 7**).

Research Findings

We answer our guiding research question (*Do we find empirical support for the microfoundations of introducing competition and choice within liberalised public infrastructure services?*) by summarising the evidence that was obtained in this study. This was done through an empirical assessment of the two microfoundations of competition and choice that were identified in the first chapter of this dissertation, the Homo Oeconomicus assumption and the homogeneous rationality assumption.

Homo Oeconomicus assumption

In chapter 3, we have used the theory of choice overload and applied it to citizens' reactions to failing public infrastructure services. Choice-overload states that increasing the number of alternatives reduces people's motivation to choose. We extended and tested the theory of choice-overload by investigating whether or not increasing the number of providers of electricity services has detrimental effects on peoples' motivation to switch their provider after a service failure. We conducted a survey experiment were the number of service providers in a service failure scenario was randomly varied. Results showed that increasing provider choice significantly reduces people's likelihood of switching away from a poor performing provider. These findings also held when replicating the experiment with an independent online sample. Thus the results from chapter 3 indicate that increasing provider choice in public infrastructure markets causally influences people's motivation to not switch away from poor performing public services. Hence our first research question (*Do citizens in their role as customers become more likely to switch away from their current public service provider after experiencing a service failure and when choice is increased?)* can be answered with a no.

Homogenous rationality assumption

To investigate the homogenous rationality assumption we asked whether different layers of society – particular those who are better versus less well educated – differ in their abilities to send markets signals to service providers in public infrastructure

markets were competition and choice have been implemented to varying degrees. Chapter 4 investigated the equality in citizens' choice behaviour (i.e. switching to another service provider) with regard to liberalised public infrastructure services across twenty-five countries of the European Union. For our analysis we combined individual-level self-reported data about citizens' market behaviour with country-level information about the market characteristics of European public infrastructure services using hierarchical modelling techniques. Our findings suggest that potentially vulnerable and non-vulnerable groups of citizens do not send market signals in the very same manner under different degrees of choice, as assumed by the homogenous rationality assumption. Indeed, our analysis revealed that the gap between lesser and better educated public service users, in terms of actual switching behaviour, widens once a considerable degree of service liberalisation has been achieved. In other words, less well educated people are less likely to have switched suppliers in the past two years in service markets where there exist a larger number of service providers. However, this choice-gap is apparent only once a certain threshold of choice is reached (more than eight providers) and was observed in the strongly liberalised mobile telephony sector, but not for less competitive services such as fixed telephony. Therefore, research question 2 (*Do different layers of society (particular those who are better versus less well educated) differ in their abilities to send markets signals to service providers in public service markets were choice has been implemented to varying degrees*?) can be partly answered with a yes.

In addition, chapter 5 has looked at whether citizens' complaint behaviour differs according to people's level of education. While citizens who are potentially vulnerable as customers are indeed less likely to have complained about any aspect of their used service in the past two years, we do not find evidence that this gap is increasing over time. Results from this chapter add to our second research question by illustrating that it is not only that the less well educated citizens are less likely to switch among service providers, but they are also less likely to complain about the services they receive. Given previous works that have illustrated that especially this group of service users is least satisfied with the services they receive (for an overview see Clifton et al., 2012), this is an important finding. It may point to the fact that potentially vulnerable citizens as customers are indeed locked-in with poor performing services. However, whether this relates to the process of service liberalisation was to be tested.

Therefore, in chapter 6 we have investigated whether a lower frequency of market signalling that potentially vulnerable service users exhibit – as found in chapters 4 and 5 – relates to the welfare gains, or losses, they get when being potentially locked-in. While it is difficult to come up with an appropriate measure of individual

welfare gains (a one Euro increase in welfare may have a different value across societal layers), we approach it from a subjective perspective by examining peoples' self-perceived affordability evaluations of the services they currently use. To address our research question, we performed a multilevel analysis of data combining individual affordability evaluations with country-level indicators about market structures of public infrastructure services across the EU-25 member states. In doing so, we related peoples' levels of self-perceived affordability with two measures of competition and choice: (1) the competitive market structure (whether choice and competition is available) – that is choice from the supply side – and (2) the competitive market functioning (whether choice is taken up by consumers) – that is choice from the demand-side. Our empirical findings suggest that less educated citizen-customers tend to experience lower levels of affordability of the services under analysis, when compared to those who are better educated. However, this gap is not significantly affected by markets' competitive structure. Instead, where this competitive structure is translated into a higher frequency of switching, the affordability gap between different socio-educational layers tends to be smaller and eventually disappears. Put simply, effective choice can reduce inequalities. Therefore we can answer our third research question (*Are there differences in individual welfare between better and less well educated citizens in public service markets were competition and choice have been implemented to varying degrees*?) with a yes. However, we have to note that the direction of the relationship is of the opposite direction as theoretically expected.

Cross-national measurement

Towards the end of this dissertation we have examined how to study the microfoundations of competition and choice through cross-national survey data of citizen satisfaction with public infrastructure services, while simultaneously accounting for respondents' heterogeneous response behaviours across countries (*research question 4*). Within this methodological contribution, we have shown that for conducting meaningful cross-national analyses one needs to consider the cross-national equivalence of used survey measures. Estimating inferential models from comparative data without taking into account the possibility of measurement non-equivalence can lead to spurious results and misleading conclusions. To illustrate this, we have estimated the biases one may get when pursuing the default approach of simply assuming the equivalence of measurement constructs by examining (1) citizen satisfaction with electricity services, and (2) respondents' trust in public institutions. We have done so by using two innovative measurement techniques that have previously received little or no attention in the public administration literature, namely multiple groups confirmatory factor analysis, and multilevel mixture item response theory. Thus this chapter contributes to this dissertation by

sketching out a methodological roadmap of how to study the microfoundations of competition and choice in future studies using cross-national public opinion data.

Conclusions

Our empirical results regarding the Homo Oeconomicus assumption provide evidence to suggest that choice-overload limits people's ability to respond to organisational failure because of the cognitive biases they face. Yet the ability of citizens to send market signals to poor performing service providers is one of the key assumptions put forward by theories of public service competition. In response, it is assumed that service providers would adjust their services to more closely match citizen's demands and preferences. But given the evidence of a choice-overload effect, the extent to which a match between citizen's preferences and demands and the offered services will be achieved can be questioned. In addition, the results of chapter 3 highlight that increasing provider choice in public service markets can potentially result in consumer inertia; this means that public service users could become locked-in to a suboptimal provider simply due to an overload of choices. Indeed, these findings stand in stark contrast with neo-classical economic thought of individuals acting as rational utility maximisers.

If we consider the theoretical rationale that in competitive markets providers adjust their goods and services in response to their customers' market signals, then our findings have severe implications for theories of public service delivery under competition. We have shown that potentially vulnerable service users in strongly marketised public service markets are less likely to switch, sending less market signals to providers then their better-off counterparts. Also they are less likely to complain about the services they receive. As a result, providers have fewer incentives to adjust the delivery of their services in accordance to their customers' needs and demands. This in turn may result in a decline in public service performance for those services used by potentially vulnerable citizens. In such a scenario, public infrastructure services may disproportionally benefit better-off service users. Indeed, this is what Clifton et al. (2011a; 2011b) and Clifton, Díaz-Fuentes and Fernández-Gutiérrez (2014) have found in a series of studies (see also Fernández-Gutiérrez, 2011). Ultimately this may contributes to citizens potentially vulnerable as customers being locked-in with poor performing providers. This would represent a serious threat to the equality criterion of the European social model that underlies the establishment of services of general economic interest (see Clifton, Comin and Díaz-Fuentes, 2005; Héritier, 2001; 2002; Prosser, 2005).

However, the evidence we have obtained in this dissertation suggests that in countries where public infrastructure services' market competition is translated into a higher levels of national switching rates, the observed affordability gap between different socio-educational groups largely decreases. In other words, inequalities can be reduced and eventually disappear when more people switch within national markets. This suggests that public infrastructure providers more closely match their service offers – in terms of the prices – with citizen-customers' demands in markets with high national switching rates. We have applied a theory of marginal consumers to explain this relationship, by suggesting that once a critical mass of citizens in their role as customers switch among service providers, thereby creating market pressures from the demand-side, this creates positive externalities to all service users, including those who are potentially vulnerable in the marketplace. Thus, it seems, that competitive markets work after all, but only once a competitive market structure is translated into higher switching rates.

In our dissertation we have shown that some, but not all, citizens deviate from exhibiting a rational market behaviour – they are bounded rational. Moreover, we show that there exist important differences between societal layers when it comes to sending market signals to providers. And despite being least satisfied, the potentially vulnerable are indeed less likely to complain to, or switch between providers of public infrastructure services. We have argued that this is so because of their limited capabilities in processing information and subsequent risk-assessments. We furthermore show that the switching-gap between these societal groups is larger in competitive markets where there exist greater opportunities to choose from. Or put simply, choice may increase inequalities in people's switching behaviour. However, our empirical analysis of whether these inequalities translate into welfare losses suggests that the right mix between switchers and non-switchers can potentially wipe-out inequalities between potentially vulnerable citizens as customers and their better-off counterparts. Thus, our dissertation puts the focus away from market regulation policies and puts the spotlight on enhancing customer empowerment and decreasing administrative burdens for exercising choice. Or put simply, for determining, and subsequently improving, the outcomes of liberalisation reforms, greater attention needs to be placed on the demand-side of public services. We hope this contribution has helped to shift a greater focus to the important role individual-level behaviours plays in determining aggregate level social outcomes of public services.

SAMENVATTING

Onderzoeksvragen

In de afgelopen decennia werd de nutssector opengesteld voor concurrentie tussen meerdere aanbieders (Conway, Janod en Nicoletti, 2005). Als gevolg hiervan moeten organisaties uit de publieke sector met andere aanbieders concurreren om te kunnen voortbestaan en kunnen zij niet langer als monopolist op hun lauweren rusten. Dat was niet zonder reden. Ingegeven door de totstandkoming van de Europese interne markt en door de publiekekeuzetheorie, werd aangenomen dat aanbieders bij meer concurrentie op de markt zouden worden gestimuleerd om betere diensten voor lagere prijzen te leveren om te kunnen voortbestaan. Met andere woorden: na succesvolle liberaliseringen zouden aanbieders moeten concurreren om klanten te winnen. Deze redenering impliceert dat burgers in de praktijk de keuze zouden hebben om bij dienstenaanbieders te vertrekken (Hirschman, 1970, pp. 21-25). Een belangrijke eigenschap bij de levering van publieke nutsvoorzieningen is immers dat de klassieke 'exit'-optie, waarbij in het geheel geen gebruik meer wordt gemaakt van de betreffende dienst, vaak niet haalbaar, te lastig of extreem duur is (zie Clifton et al., 2012). Zo zouden afnemers die besluiten om bij een leverancier van gas of elektriciteit te vertrekken waarschijnlijk met ernstige gevolgen voor hun individuele welzijn te maken krijgen, vooral in de winter. Hier wordt het begrip 'keuze' van belang, dat wil zeggen het overstappen naar andere (publieke of private) aanbieders als subvorm van de 'exit'-optie (Dowding en John, 2012). Dus door over te stappen, of middels klachten impliciet te dreigen dat te gaan doen (dit noemt Hirschman [1970] het 'voice'-signaal), geven burgers marktsignalen af aan aanbieders. Dat zou dan werken als een prikkel om afnemers meer waar voor hun geld te leveren, zodat bestaande klanten kunnen worden behouden en nieuwe kunnen worden geworven. Daarom wordt aangenomen dat op de lange termijn een evenwicht zal ontstaan tussen vraag en voorkeuren van burgers enerzijds en prijs en kwaliteit van de aangeboden diensten anderzijds.

In deze neoklassieke economische zienswijze op publieke dienstverlening met concurrentie wordt doorgaans aangenomen dat het introduceren van concurrentie in de publieke dienstverlening en een toename van het aantal dienstenaanbieders waaruit afnemers een keus kunnen maken uiteindelijk zal leiden tot optimale verdeling van de beschikbare middelen. Met andere woorden: de publieke dienstverlening zou beter en goedkoper worden. Dit is gebaseerd op een significante reeks aannames op microniveau, de microgrondslagen van concurrentie en keuze.

In de theoretische beschouwingen over de invoering van concurrentie en keuze in de publieke dienstverlening worden belangrijke aannames gedaan over motivatie en gedrag op individueel niveau, die met name voortkomen uit de gedachte dat

mensen rationeel en berekenend op zoek zijn naar de beste voorzieningen (zie ook Stoker en Moseley, 2010). Hervormingen op macroniveau, waarbij in de publieke dienstverlening marktwerking wordt ingevoerd, zijn dan ook opgebouwd rond het concept van de *Homo Oeconomicus*. Voor de vraag of deze microgrondslagen van concurrentie en keuze wel juist zijn, is er echter slechts in beperkte mate aandacht (dit is derhalve empirisch nog veel minder concreet aangetoond). In dit onderzoek hebben we gekeken naar twee belangrijke aannames op microniveau die ten grondslag liggen aan de invoering van concurrentie en keuze in de publieke dienstverlening: In de eerste plaats zouden **burgers zich op geliberaliseerde markten voor publieke diensten als rationele klanten gedragen** (*de Homo Oeconomicus-aanname*). Wanneer burgers meer keuzemogelijkheden hebben, zo wordt aangenomen, zouden zij het dienstenaanbod zoeken en kiezen dat het beste aansluit bij hun vraag en behoeften (Europese Commissie, 2004; Stone, 2005). Met andere woorden: zij kiezen weloverwogen de aanbieder die hun persoonlijk de beste voorzieningen kan bieden. In de tweede plaats zouden **alle burgers zich als klant even rationeel gedragen.** Dit noemen we de *homogene-rationaliteitsaanname*. Dit zou betekenen dat alle burgers in hun rol van klant de beslissingen nemen die hun persoonlijke optimum zoveel mogelijk benaderen en dus hun welvaart homogeen doen toenemen. Anders gezegd: er wordt aangenomen dat alle burgers in hun rol van klant even goed in staat zijn om in te spelen op de kansen die worden geboden door de toename van concurrentie en keuze. Alle groepen burgers gedragen zich als klant even rationeel. Bij de liberaliseringen is dus niet alleen aangenomen dat alle groepen burgers even rationeel handelen, maar ook dat onder die burgers de toename van individuele welvaart gelijkelijk zou zijn verdeeld (Europese Commissie, 2004; zie ook Clifton et al., 2011a).

In dit proefschrift hebben wij de beperkingen onderzocht van rationele beschrijvingen van de manier waarop burgers zich gedragen op markten voor publieke dienstverlening. Aan de hand van gedragsonderzoek naar individuele besluitvorming bieden wij een andere kijk op de manier waarop burgers zich daadwerkelijk gedragen op geliberaliseerde markten voor publieke dienstverlening en toetsen wij dit empirisch. Concreet hebben wij onderzocht of burgers inderdaad rationele beslissingen nemen in de nutsector waar concurrentie en keuze zijn ingevoerd. Daarbij hebben wij een reeks onderzoeksvragen getoetst die van groot theoretisch en praktisch belang zijn voor de manier waarop de huidige publieke dienstverlening is georganiseerd. Eerst onderzochten wij de Homo Oeconomicus-aanname door de vraag te stellen of burgers, als zij meer keuze hebben, naar andere aanbieders overstappen na slechte ervaringen met de dienstverlening (**hoofdstuk 3**). Vervolgens onderzochten wij de homogene rationaliteitsaanname door te kijken of verschillende lagen van de samenleving (met name beter opgeleide in vergelijking

met minder goed opgeleide) ook verschillen in hun vermogen om van dienstverlener te veranderen op markten voor publieke dienstverlening waar in uiteenlopende mate keuzemogelijkheden zijn ingevoerd, en of we vergelijkbare patronen zien voor hun klachtengedrag (**hoofdstuk 4 en 5**). Daarna hebben we gekeken of verschillen in de manier waarop burgers zich op geliberaliseerde markten voor publieke dienstverlening als klant gedragen ook van invloed kunnen zijn op de mate waarin zij van deze marktordeningen 'profiteren', en of de diverse lagen van de samenleving hierin verschillen. Wij hebben dus onderzocht of ongelijkheden tussen hoger en lager opgeleide afnemers groter of kleiner zijn op markten waarin keuze en concurrentie in uiteenlopende mate zijn ingevoerd (**hoofdstuk 6**). Aan het eind van het onderzoek hebben we een methodologische visie opgenomen over de manier waarop de microgrondslagen van concurrentie en keuze moeten worden onderzocht aan de hand van onderzoeksgegevens over de tevredenheid van burgers met publieke nutsvoorzieningen. We hebben tegelijkertijd het heterogene responsgedrag verklaard van respondenten in verschillende landen (**hoofdstuk 7**).

Onderzoeksbevindingen

Wij beantwoorden onze centrale onderzoeksvraag (*Vinden we empirische ondersteuning voor de microgrondslagen van de invoering van concurrentie en keuze binnen geliberaliseerde, publieke nutsvoorzieningen?*) aan de hand van een overzicht van de in dit onderzoek verzamelde resultaten. Wij hebben dit gedaan door middel van een empirische beoordeling van de beide microgrondslagen voor concurrentie en keuze die in het eerste hoofdstuk van dit proefschrift zijn vastgesteld, te weten de Homo Oeconomicus-aanname en de homogene rationaliteitsaanname.

Homo Oeconomicus-aanname

In hoofdstuk 3 gebruikten we de theorie van overdaad aan keuze ('choice overload') en pasten wij deze toe op de reacties van burgers op falende publieke nutsvoorzieningen. Volgens deze theorie zijn mensen minder gemotiveerd om te kiezen als zij meer keuzemogelijkheden krijgen. We hebben de theorie van overdaad aan keuze uitgebreid en getoetst door te onderzoeken of een toename van het aantal aanbieders van elektriciteit al dan niet nadelig werkt op de motivatie van mensen om bij falende dienstverlening naar een andere aanbieder over te stappen. Wij deden een onderzoeksexperiment waarbij het aantal aanbieders willekeurig varieerde bij tekortkomingen in de dienstverlening. De kans dat mensen bij een slecht presterende aanbieder weggaan, bleek significant af te nemen naarmate zij uit meer aanbieders konden kiezen. Deze bevindingen hielden ook stand toen het experiment werd herhaald met een onafhankelijke online steekproef. De resultaten

uit hoofdstuk 3 wijzen er dus op dat meer keuze uit aanbieders in de nutssector de motivatie van mensen om niet over te stappen wanneer de publieke dienstverlening slecht is causaal beïnvloedt. Onze eerste onderzoeksvraag (*Gaan burgers in hun rol van klant eerder bij hun huidige aanbieder van publieke diensten weg wanneer de dienstverlening tekortschiet en zij meer keuzemogelijkheden hebben?)* kan dus met 'nee' worden beantwoord.

Homogene-rationaliteitsaanname

Om deze aanname te onderzoeken, stelden wij de vraag of verschillende lagen van de samenleving – met name beter opgeleide in vergelijking met minder goed opgeleide – ook verschillen in hun vermogen om marktsignalen af te geven aan aanbieders in de nutssector waar concurrentie en keuze in uiteenlopende mate zijn ingevoerd. Hoofdstuk 4 onderzocht de gelijkheid in het keuzegedrag van burgers (d.w.z. het overstappen naar een andere aanbieder) voor geliberaliseerde, publieke nutsvoorzieningen in 25 landen van de Europese Unie. Voor onze analyse combineerden we individuele zelf gerapporteerde data over het marktgedrag van burgers met landspecifieke informatie over de marktkenmerken van Europese, publieke nutsvoorzieningen. Daarbij maakten we gebruik van hiërarchische modelleertechnieken. Onze bevindingen wijzen erop dat potentieel kwetsbare en niet-kwetsbare groepen burgers niet op dezelfde manier marktsignalen afgeven wanneer hun keuzemogelijkheden verschillen, zoals wordt verondersteld in de homogene-rationaliteitsaanname. Uit onze analyse bleek namelijk dat het verschil tussen lager en hoger opgeleide gebruikers van publieke diensten, waar het gaat om feitelijk overstapgedrag, groter wordt wanneer een aanzienlijke mate van liberalisering van de dienstverlening tot stand is gekomen. Met andere woorden: lager opgeleiden zullen in de afgelopen twee jaar waarschijnlijk minder makkelijk van leverancier zijn veranderd op dienstenmarkten waar een groter aantal aanbieders aanwezig is. Dit keuzeverschil is echter alleen zichtbaar nadat een bepaalde drempel aan keuzemogelijkheden is bereikt (meer dan acht aanbieders) en werd waargenomen in de sterk geliberaliseerde mobiele telefoniesector, maar niet bij minder concurrerende diensten, zoals vaste telefonie. Derhalve kan onderzoeksvraag 2 (*Bestaat er tussen verschillende lagen van de samenleving (met name hoger opgeleide lagen in vergelijking met lager opgeleide) een verschil in hun vermogens om marktsignalen af te geven aan aanbieders op markten voor publieke dienstverlening waar in uiteenlopende mate keuze is ingevoerd*?) ten dele met 'ja' worden beantwoord.

Vervolgens is in hoofdstuk 5 gekeken of het klachtgedrag van burgers verschilt naargelang het opleidingsniveau van mensen. Hoewel burgers die als klant potentieel kwetsbaar zijn inderdaad minder snel klagen over aspecten van de dienst die zij in de afgelopen twee jaar hebben afgenomen, hebben wij niet kunnen aantonen

dat dit verschil in de loop der tijd groter wordt. De resultaten uit dit hoofdstuk zijn relevant voor onze tweede onderzoeksvraag omdat zij niet alleen laten zien dat lager opgeleide burgers minder snel naar een andere aanbieder zullen overstappen, maar ook dat zij minder snel zullen klagen over de diensten die zij geleverd krijgen. Gelet op eerder onderzoek waaruit bleek dat met name deze groep afnemers van diensten het minst tevreden is over de diensten die zij afnemen (zie voor een overzicht Clifton et al., 2012), is dit een belangrijke uitkomst. Mogelijk wijst dit erop dat potentieel kwetsbare burgers als klant inderdaad vastzitten aan slecht presterende diensten. Wel moest worden onderzocht of dit ook verband houdt met het liberaliseringsproces van diensten.

Daarom hebben wij in hoofdstuk 6 onderzocht of het feit dat potentieel kwetsbare afnemers van diensten minder vaak marktsignalen afgeven (zoals we zagen in de hoofdstukken 4 en 5) verband houdt met hun toegenomen (of afgenomen) welvaart wanneer zij mogelijk vastzitten. Hoewel het moeilijk is om een individuele toename van welvaart goed te meten (als deze met één euro toeneemt, heeft dat niet in alle lagen van de samenleving dezelfde waarde), benaderen wij dit vanuit een subjectief perspectief en kijken we wat mensen zelf vinden van de betaalbaarheid van de diensten die zij momenteel gebruiken. Om onze onderzoeksvraag te beantwoorden, maakten wij een gegevensanalyse op meerdere niveaus, waarin individuele beoordelingen van de betaalbaarheid werden gecombineerd met landspecifieke indicatoren over marktstructuren voor publieke nutsvoorzieningen in de lidstaten van de EU-25. Daarbij koppelden wij de eigen mening van mensen over de betaalbaarheid aan twee metingen van concurrentie en keuze: (1) de structuur van een markt met concurrentie (de vraag of keuze en concurrentie aanwezig zijn) – dus keuze aan de aanbodzijde – en (2) de werking van een markt met concurrentie (de vraag of consumenten van de keuzemogelijkheden gebruik maken) – dus keuze aan de vraagzijde. Onze empirische bevindingen wijzen erop dat lager opgeleide burgers in vergelijking met de hoger opgeleiden de onderzochte diensten vaak minder betaalbaar blijken te vinden. Dit verschil wordt echter niet significant beïnvloed door de concurrentiestructuur van markten. Wanneer er binnen deze concurrentiestructuur vaker wordt overgestapt, zal het verschil in betaalbaarheid tussen maatschappelijke lagen met verschillende opleidingsniveaus doorgaans kleiner zijn en uiteindelijk verdwijnen. Eenvoudig gezegd: ongelijkheden kunnen worden verkleind door effectieve keuzemogelijkheden. Onze derde onderzoeksvraag (*Zijn er verschillen in individuele welvaart tussen hoger en lager opgeleide burgers op markten voor publieke dienstverlening waarop concurrentie en keuze in verschillende mate zijn ingevoerd*?) kunnen wij daarom met 'ja' beantwoorden. Wel moet echter worden opgemerkt dat de richting van de relatie tegengesteld is aan de theoretisch verwachte richting.

Cross-nationale meting

Aan het eind van dit proefschrift hebben we bekeken op welke manier de microgrondslagen van concurrentie en keuze moeten worden onderzocht aan de hand van cross-nationale onderzoeksgegevens over de tevredenheid van burgers met publieke nutsvoorzieningen, waarbij tegelijkertijd aandacht wordt geschonken aan het heterogene responsgedrag van de respondenten in de verschillende landen (*onderzoeksvraag 4*). In deze methodologische bijdrage hebben wij laten zien dat voor het uitvoeren van zinvolle cross-nationale analyses gekeken moet worden naar de cross-nationale gelijkwaardigheid van de gebruikte onderzoeksmetingen. Wanneer bij het beoordelen van inductieve modellen uit vergelijkende data geen rekening wordt gehouden met de mogelijkheid dat metingen niet gelijkwaardig zijn, kan dat tot onjuiste resultaten en misleidende conclusies leiden. Om dit te illustreren hebben we de vertekeningen opgemeten die kunnen ontstaan wanneer de standaardbenadering wordt gevolgd en eenvoudigweg wordt aangenomen dat meetconstructen equivalent zijn door (1) de tevredenheid van burgers met leveranciers van elektriciteit en (2) het vertrouwen van respondenten in overheidsinstanties te onderzoeken. Wij hebben daarbij gebruik gemaakt van twee innovatieve meettechnieken waaraan voorheen in de literatuur over overheidsdiensten weinig tot geen aandacht werd geschonken, namelijk een bevestigende meergroepen-factoranalyse en een multiniveau-mix itemresponstheorie. De bijdrage van dit hoofdstuk aan dit proefschrift bestaat dus uit het schetsen van een methodologische routekaart voor de wijze waarop de microgrondslagen van concurrentie en keuze moeten worden onderzocht in toekomstige studies met gebruikmaking van cross-nationale gegevens inzake de publieke opinie.

Conclusies

Op grond van onze empirische resultaten met betrekking tot de Homo Oeconomicus-aanname mag worden verondersteld dat mensen bij een overdaad aan keuzemogelijkheden minder goed in staat zijn om op tekortkomingen van organisaties te reageren vanwege de cognitieve bias waarmee zij te maken hebben. Toch is het vermogen van burgers om marktsignalen af te geven aan slecht presterende aanbieders een van de belangrijkste aannames uit theorieën over concurrentie in de openbare dienstverlening. In reactie daarop wordt aangenomen dat aanbieders hun diensten aanpassen om deze beter af te stemmen op vraag en voorkeuren van de burger. Vanwege het aangetoonde gevolg van een overdaad aan keuzemogelijkheden mogen echter vraagtekens worden geplaatst bij de mate waarin vraag en voorkeuren van de burger en de aangeboden diensten op elkaar kunnen worden afgestemd. Daarnaast maken de resultaten van hoofdstuk 3 duidelijk dat meer keuze uit aan-

bieders op markten voor publieke diensten ertoe kan leiden dat de consument blijft zitten waar hij zit. Dat betekent dat afnemers van publieke nutsvoorzieningen vast zouden kunnen komen te zitten aan een niet optimale aanbieder, eenvoudigweg omdat ze teveel keuzemogelijkheden hebben. Deze bevindingen staan inderdaad in sterk contrast met de neoklassieke economische gedachte dat mensen rationeel zoeken naar een zo groot mogelijk voordeel.

Als we kijken naar de theoretische rationale dat aanbieders op concurrerende markten hun goederen en diensten aanpassen aan de marktsignalen van hun klanten, dan hebben onze bevindingen ernstige implicaties voor theorieën over publieke dienstverlening op een voor concurrentie opengestelde markt. Wij hebben laten zien dat potentieel kwetsbare gebruikers van diensten op dergelijke markten voor publieke dienstverlening minder tot overstappen geneigd zijn en minder marktsignalen aan aanbieders afgeven dan anderen die in een betere positie verkeren. Ook zullen zij minder snel klagen over de diensten die hun worden geleverd. Hierdoor worden aanbieders minder gestimuleerd om hun diensten aan te passen aan vraag en behoeften van hun klanten. Dat kan vervolgens leiden tot een achteruitgang van de prestaties van de publieke dienstverlening in het geval van diensten die worden afgenomen door potentieel kwetsbare burgers. In een dergelijk scenario kunnen publieke nutsvoorzieningen onevenredig voordelig zijn voor beter gesitueerde afnemers van die diensten. Dit is namelijk wat Clifton et al.(2011a; 2011b) en Clifton, Díaz-Fuentes en Fernández-Gutiérrez (2014) hebben geconcludeerd in een reeks studies (zie ook Fernández-Gutiérrez, 2011). Uiteindelijk kan dit er mede toe bijdragen dat potentieel kwetsbare burgers als klant vast komen te zitten aan slecht presterende aanbieders. Dat zou een ernstige bedreiging zijn van het gelijkheidscriterium van het Europees sociaal model dat ten grondslag ligt aan de totstandkoming van diensten van algemeen economisch belang (zie Clifton, Comin en Díaz-Fuentes, 2005; Héritier, 2001; 2002; Prosser, 2005).

De uitkomsten van dit proefschrift lijken er echter op te wijzen dat in landen waar concurrentie in de nutssector zich voor die landen vertaalt in grotere aantallen overstappers, het waargenomen verschil in betaalbaarheid tussen maatschappelijke groepen met verschillende opleidingsniveaus sterk afneemt. Met andere woorden: ongelijkheden kunnen worden verkleind en verdwijnen uiteindelijk wanneer op nationale markten meer mensen overstappen. Dit wijst erop dat aanbieders van publieke nutsvoorzieningen hun dienstenaanbod (waar het gaat om prijzen) beter afstemmen op de vraag van burgers op markten in landen met veel overstappers. Wij hebben een marginale consumententheorie toegepast om deze relatie uit te leggen, door te suggereren dat wanneer eenmaal een kritische massa van burgers in hun rol als klant naar andere aanbieders van diensten overstappen en daarmee

marktdruk aan de vraagzijde creëren, dit positieve externe effecten heeft voor alle afnemers van diensten, waaronder degenen die op de markt potentieel kwetsbaar zijn. Zo lijkt het dus dat concurrerende markten uiteindelijk wel werken, maar alleen nadat er op markten die voor concurrentie zijn opengesteld meer overstappers zijn gekomen.

In ons proefschrift hebben wij laten zien dat sommige (maar niet alle) burgers geen rationeel marktgedrag vertonen – zij zijn rationeel gebonden. Bovendien laten we zien dat er tussen de lagen in de maatschappij belangrijke verschillen bestaan waar het gaat om het afgeven van marktsignalen aan aanbieders. En hoewel zij het minst tevreden zijn, zullen de potentieel kwetsbare mensen inderdaad minder snel klagen of overstappen naar andere aanbieders van publieke nutsvoorzieningen. Wij hebben betoogd dat dit het geval is omdat zij beperkt in staat zijn om informatie te verwerken en vervolgens een inschatting van de risico's te maken. Ook laten wij zien dat de verschillen in aantallen overstappers tussen deze groepen in de samenleving groter zijn op voor concurrentie opengestelde markten met meer keuzemogelijkheden. Eenvoudig gezegd: keuzemogelijkheden kunnen ongelijkheden in het overstapgedrag van mensen vergroten. Onze empirische analyse van de vraag of deze ongelijkheden zich in minder welvaart vertalen wijst er echter op dat de juiste samenstelling tussen overstappers en blijvers potentieel de ongelijkheden kan doen verdwijnen tussen potentieel kwetsbare burgers als klant en beter gesitueerde anderen. Ons proefschrift verschuift de aandacht dus van marktreguleringsbeleid naar verbetering van de mondigheid van klanten en verkleining van de administratieve lasten bij het uitoefenen van keuzevrijheid. Anders gezegd: om vast te stellen tot welke uitkomsten liberaliseringen hebben geleid en deze uitkomsten vervolgens te verbeteren, moet gewoon meer aandacht worden geschonken aan de vraagzijde van publieke diensten. Wij hopen dat er mede door deze bijdrage meer aandacht wordt geschonken aan de belangrijke rol van individueel gedrag bij het bepalen van de maatschappelijke gevolgen op macroniveau van publieke diensten.

ABOUT THE AUTHOR

Sebastian Jilke studied Social Management (2004-2007) at the University of Applied Sciences of the German Red Cross in Göttingen (Germany). In 2010, he obtained his Masters degree in Public Management and Governance (including a minor in Corporate Management and Economics) at Zeppelin University in Friedrichshafen (Germany). During his MA studies, he specialised in public sector reforms in Sub-Saharan Africa, and halfway through his studies he moved to Addis Ababa (Ethiopia) to gain some practical work experiences. His master thesis "Equal Access to Basic Services in African LDCs" was awarded the ZU best master thesis award in 2010. During that time he already worked as a consultant for the World Bank in Ethiopia.

In December 2010, he started as a PhD student in Public Administration at Erasmus University Rotterdam. During that time he was the administrative manager of the FP7 funded project COCOPS "Coordinating for Cohesion in the Public Sector of the Future", where he also was involved in various work packages. He also initiated and convened the Erasmus University Department of Public Administration Seminar Series in Research Methodology (2012-2014).

Currently, Sebastian works as a postdoctoral researcher in Public Administration at Erasmus University Rotterdam. In doing so, his research focuses on the politics and psychological underpinnings of citizen's attitudes and behaviors towards public services. He is applying a behavioural perspective to questions that revolve around the possible tensions between public service delivery reforms and democracy. Thus his academic work lies at the intersection of public administration, political science, behavioural economics and social psychology – or simply: *behavioural public administration*. His secondary research interest is of methodological nature, especially in the areas of measurement and (quasi-)experimental techniques.

Sebastian is also co-chairing the IIAS study group on public attitudes and trust, and regularly organises panels and workshops at key conferences within public administration. He has published widely in leading journals in public administration and political science, including – among others – the European Journal of Political Research, Public Administration, Public Administration Review and the International Public Management Journal. Moreover, he is the organising co-editor for a symposium at the Public Administration Review on the use of experimentation in public administration research. Moreover, together with coauthors, he is currently developing an edited book on the use of experiments.

Sebastian Jilke
jilke@fsw.eur.nl
www.sebastianjilke.net