The Economics of Net Neutrality

Implications of Priority Pricing in Access Networks

Zur Erlangung des akademischen Grades eines
Doktors der Wirtschaftswissenschaften

(Dr. rer. pol.)

von der Fakultät für Wirtschaftswissenschaften
des Karlsruher Institut für Technologie (KIT)

genehmigte

DISSERTATION

von

Dipl. Volksw. Lukas Wiewiorra

Tag der mündlichen Prüfung: 10.07.2012
Referent: Prof. Dr. Christof Weinhardt
Korreferent: Prof. Dr. Karl-Martin Ehrhart
Karlsruhe, 2012

Acknowledgements

Foremost I would like to thank Prof. Dr. Christof Weinhardt, who gave me the great chance to start my work at the Institute of Information Systems and Management (IISM) and who supported my studies at the DFG Graduate School Information Management and Market Engineering (IME) at the Karlsruhe Institute of Technology (KIT). He offered me the opportunity and freedom to pursue this research. I am also indebted to my co-advisor Prof. Dr. Karl-Martin Ehrhart for his encouragement and his constructive comments throughout my studies at the IME. I would particularly like to thank the other members of the committee, Prof. Dr. Rudi Studer for the interesting discussions on practical implications of my thesis and Prof. Dr. Kay Mitusch for having been so kind to chair the board of examiners.

Moreover, I would like to thank the team of IISM and the professors and students of the IME who have helped to improve this work through constructive feedback and who have created an excellent working atmosphere and ensured a vivid social life over the last years. I would also like to thank Sascha Schweitzer and Philip Köhler, who have been my roommates within the last years and both helped to make the process of creating this work an unforgettable journey. In particular I want to thank Dr. Jan Krämer who has been not only an outstanding mentor, motivator and colleague since the day I started my thesis, but also became a dear friend.

Furthermore, I want to thank Christoph Helbach for encouraging me in my initial idea to pursue a doctoral degree, Stephan Meyer for his perennial supporting accompaniment and Rebekka Freitag-Li for proof-reading this thesis and giving valuable comments that improved this work a lot.

Finally, I would like to thank those people without whom this work would not have been possible. I would like to express my thankfulness to my parents Monika and Ulrich Wiewiorra for their unconditional support in everything I did and their love throughout my whole life, my brother Jeremias Wiewiorra for reminding me of the humorous side of life and Bettina for her love, support, forbearance and her caring encouragement.
I owe them more than my deepest gratitude.

Lukas Wiewiorra

Preface

"The Internet is the first thing that humanity has built that humanity doesn't understand, the largest experiment in anarchy that we have ever had."

Eric Schmidt (Chairman of Google)

OVER the last nine years a world wide debate about the future of the Internet has gained momentum. Many different opinions, arguments and scenarios have been subsumed under the umbrella of the term 'Net Neutrality (NN)'. A multitude of definitions and theories about this vague concept exists. NN is a politically heavily debated matter in many different countries. Over the recent years the focus of the NN debate has shifted towards a discussion about the effect of prioritization of data in the access networks of Internet Service Providers (ISPs).

This thesis systematically analyzes NN from an economic point of view. To this end a framework is developed which helps to structure the NN debate. Furthermore, the introduction of prioritization is studied by analyzing potential effects of Quality of Service (QoS) on Content and Service Providers (CSPs) and Internet Users (IUs).

The structure of this thesis is composed of six chapters, as shown in Figure 1. Chapter one gives an introduction to the history of the Internet and the concept of NN. Moreover, different definitions of NN are presented and the relevant economic relations in the Internet ecosystem are exemplified.
Chapter two presents a so called 'Non Net Neutrality (NNN)' framework. This framework allows to structure the NN debate along two basic dimensions and to group the multitude of scientific approaches that exist to analyze deviations from the 'status quo'. The chapter concludes with a policy decision guideline which was created in an effort to structure and summarize the identified gains, threats and possible remedies of each NNN scenario according to the dimensions of the aforementioned framework.
Chapter three deals with the CSP tiering scenario of a possible QoS tiering network regime and presents a theoretical economic model that allows to further analyze specific aspects of

the NN debate.
Chapter four deals with the IU tiering scenario of a possible QoS tiering network regime and presents an empirical analysis based on survey data of 1035 Internet access customers to analyze the perception of QoS and the profitability of IU tiering.
Chapter five presents an outlook on how the neutrality principle could evolve to other layers of the Internet ecosystem.
Chapter six summarizes the findings of this thesis and exemplifies the implications for the ongoing political process around NN and for the business of ISPs. The chapter concludes with suggestions for future research in the field of NN.

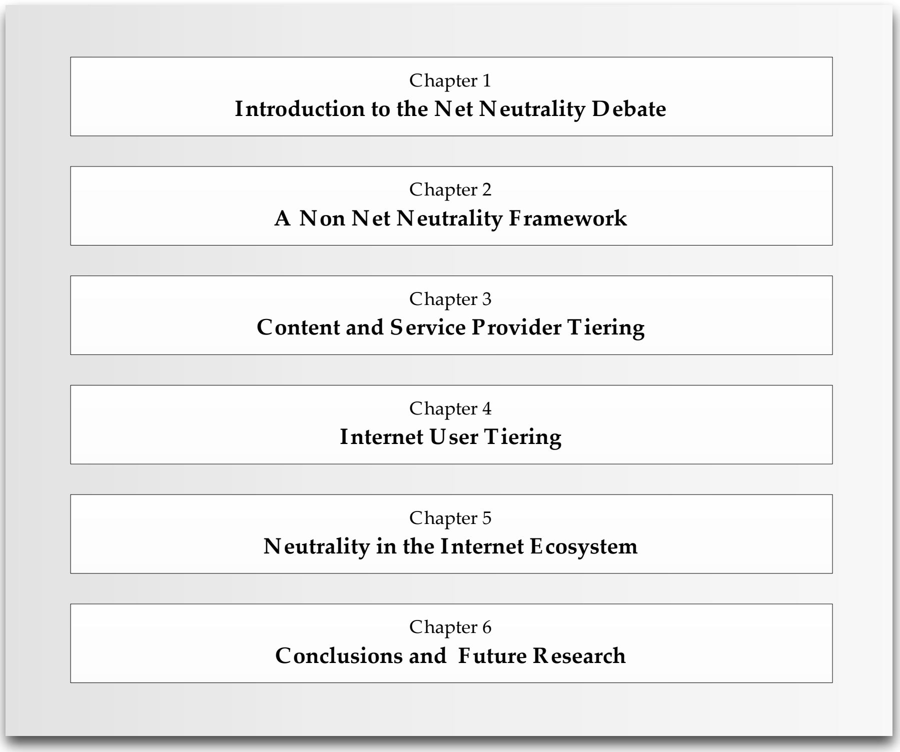

Figure 1.: Structure of this thesis

Contents

List of Figures

List of Tables

Chapter 1.

Introduction to the Net Neutrality Debate

"Supply creates it's own demand"

Say's law by John Maynard Keynes

THIS and the following chapter are intended as an introduction and progress report on the growing body of literature on the issue of NN. The chapters correspond in large parts to the paper by Krämer, Wiewiorra, and Weinhardt (2012).

The so-called NN debate centers around the potential consequences of network owners exercising additional control over the data traffic that is currently being transferred without further interference and without additional fees through their networks. This chapter gives an introduction to the history of the Internet, the concept of NN and the relevant economic relations in the Internet ecosystem. The following chapter presents a framework by which the current literature can be classified, despite that fact that the contributions come from various academic disciplines such as law, engineering and economics. With this approach, one is able to identify commonalities between seemingly different standpoints and to extract the general consensus, as well as the inherent differences between proponents and opponents of NN alike. Moreover, examples of NN regulation are discussed and a policy decision guideline is derived.

1.1. History of the Internet

Internet is the abbreviation of the term internetwork, which describes the connection between computer networks all around the world through routers and gateways on the basis of the same set of communication protocols. Starting off as a closed research network between just a few universities in 1967, the underlying network architecture still had to evolve to an open network environment.

With the transition of the existing infrastructure to the new Transmission Control Protocol / Internet Protocol (TCP/IP) at January 1, 1983, the Internet established as a communication tool for a growing number of people. The design principles for the TCP/IP were guided by two fundamental design principles: Messages are fragmented into data packets that are routed through the network autonomously (End-to-End principle) and as fast as possible (Best Effort (BE) principle). This means that intermediate nodes, so-called routers, do not differentiate packets based on their content or source. Rather, routers maintain routing tables in which they store the next node that lies on the supposedly shortest path to the packet's destination address. However, whereas each router acts autonomously along which path it sends a packet, no router has control over which path the packet is send from sender to receiver. Packets are stored in a router's queue if they arrive at a faster rate than the rate at which the router can send out packets. When the router's queue is full, the package is deleted (dropped) and must be resent by the source node. Full router queues are the main reason for congestion on the Internet. However, no matter how 'important' a data packet may be, routers would always process their queue according to the first-in-first-out principle. Especially these fundamental principles have always been a key element of the open Internet spirit and in the context of the NN debate this has become known as the *non-discrimination rule* (see, e. g., Schuett, 2010).

The first commercial ISP ('The World') started to offer public dial up access to the Internet in 1990 (Zakon, 1997). Together with the disruptive innovation of content visualization and linkage via the Hyper Text Markup Language (HTML), the so called World Wide Web (WWW) made the Internet a global success. In the dawning days of the commercial WWW in the mid nineties the infrastructure forming the growing Internet was developed without having companies in mind which have to invest heavily in bringing broadband access to every household. Since the time when access to the Internet was predominately provided by dial up connections, the average data traffic per household has severely increased with the availability of broadband and rich media content (Bauer et al., 2009). According to the 'Minnesota Internet Traffic Studies' the overall Internet traffic in the United States (US) is growing annually by approximately 50 percent (Odlyzko et al., 2012). The increase in network traffic is the consequence of the ongoing transition of the Internet to a fundamental universal access technology. Classic broadcasting media consumption is declining and content is instead consumed via the Internet. On the other hand the Internet "is evolving to permit more sophisticated forms of pricing and cost recovery, a perhaps painful requirement in this commercial world." (Leiner et al., 2011).

Today the commercial Internet ecosystem consists of several players. IU are connected to the network by their local ISP, while CSPs offer a wide range of applications and content to this mass of potential consumers. All of these actors are spread around the world and interconnect with each other via the so called backbone of the Internet, which is operated by several big network providers (Economides, 2005).

1.2. What is Net Neutrality?

With the rapid development of the Internet as an ubiquitously available platform for information, entertainment and communication, the role of network infrastructure owners shifted to an essential gatekeeper position in the information society. What caused the debate is that ISPs have signaled that they intend to use their power over the network infrastructure to generate extra revenues. In this context, opponents of these plans envision several particular deviations from the status quo which, as they say, endanger the 'openness' of the Internet that has been so fruitful for content and service innovations that have advanced the daily lives of all IUs. Consequently, the public and politicians alike are concerned about how ISPs are going to monetize access and usage of the networks in the future. The debate was particularly stimulated in 2005 after a blunt statement by Ed Whitacre, at the time the Chief Executive Officer of ATT, who said: "Now what [content providers] would like to do is use my pipes free, but I ain't going to let them do that because we have spent this capital and we have to have a return on it" (O'Connell, 2005). Similar statements have been released by major European network operators since then (Lambert, 2010; Schneibel and Farivar, 2010).

The term 'Net Neutrality' was coined by law professor Tim Wu (2003) a strong advocate of NN who has still a large influence on the direction of the ongoing debate. However, the discussion of neutrality in the Internet roots in the open access movement that was led by Lawrence Lessig (2001, p.168–175).

From this point the term NN became the figurehead of a debate that centers around the potential consequences of network owners exercising additional control over the data traffic that is currently being transferred without further interference (BE principle) and without additional fees through their networks. In this context, the meaning of 'control' is often ambiguous and can mean anything from blocking certain types of undesired or unaffiliated traffic (Wu, 2007), to termination fees (Lee and Wu, 2009), to offering differentiated services and taking measures of network management (Hahn and Wallsten, 2006).

The term Net Neutrality has no distinct definition. Each pressure group in the debate has its own construction of the term. In the US consumer rights groups are fighting fiercely against the lobbying of big telecommunications and cable companies. Their definition of NN is in particular aimed against a tiered system:

Definition 1.1 (Consumer rights group)**.** *Net Neutrality means no discrimination. Net Neutrality prevents Internet providers from speeding up or slowing down Web content based on its source, ownership or destination. (Save the Internet, 2011, FAQ)*

In their paper Hahn and Wallsten (2006) acknowledge that, “Net neutrality has no widely accepted precise definition, but[...]”

Definition 1.2 (Hahn and Wallsten)**.** *“[...]usually means that broadband service providers charge consumers only once for Internet access, do not favor one content provider over another, and do not charge content providers for sending information over broadband lines to end users.”(Hahn and Wallsten, 2006, p.1)*

The so-called father of the Internet Sir Tim Berners-Lee is more concerned about the universal access of information and the fully interconnected character of the Internet. In an interview he made a statement, that can be interpreted as his personal ‘definition’ of the NN concept:

Definition 1.3 (Sir Berners-Lee)**.** *“While we may pay for different service levels, e.g. we pay more for a higher bandwidth, the important thing about the Net is that if we both pay for a certain level of service, then we can communicate at that level no matter who we are. We pay to be able to connect to a certain bandwidth and that’s all we have to do. It’s up to our ISPs to ensure that the interconnection is done. This is how it has always been done.”(Powell, 2006, p.3).*

Eric Schmidt, at the time of the statement CEO of Google, one of the biggest CSPs in the world, explained the company’s definition of NN as follows:

Definition 1.4 (Google and Verizon)**.** *"I want to be clear what we mean by Net neutrality: What we mean is if you have one data type like video, you don't discriminate against one person's video in favor of another, but it's okay to discriminate across different types, so you could prioritize voice over video, and there is general agreement with Verizon and Google on that issue."(Fehrenbacher, 2010)*

Over the last years the question about neutrality in Internet access has grown from a mere dispute between policy makers and network owners to an academic debate about the potential pitfalls of ex-ante regulation and the long term effects on the Internet ecosystem in case of an 'laissez faire' approach. Currently, with the exception of Chile and the Netherlands, no explicit NN law exists that would forbid ISPs to manage, differentiate or price their networks as they wish.
In summary, one can aggregate the matter of NN with the following short but essential statement.

The Net Neutrality (NN) debate centers around the question whether the potential outcome of possible NNN scenarios appears so dangerous to the freedom of the Internet ecosystem that ex-ante NN regulation is necessary and appropriate.

1.3. Non Net Neutrality - Necessity or New Business Model?

ISPs are confronted with the transformation of the Internet to an universal access technology. One prominent example for this development is the company 'Netflix' in the US. Netflix offers video on demand streaming of many TV shows and movies for a monthly subscription fee. According to Sandvine (2010, p.14) already 20.6 percent of all downstream traffic on fixed access networks in North America during peak times is Netflix. In total, approximately

45 percent of downstream traffic on fixed and mobile access networks in North America is attributable to real-time entertainment (Sandvine, 2010, p.12).

In 2001, Bret Swanson coined the term ‘exaflood’. In Swanson and Gilder (2008) he draws the picture of the presumably impending flood of exabytes[1] due to increasing media consumption that would cause the congestive breakdown of the Internet.

ISPs in the fixed line business, be it former telecommunications providers or cable companies, are concerned with this rapid development of traffic flows and the resulting infrastructure costs. This is mainly because they invest into the last mile infrastructure bringing the Internet to the individual households. The largest share of the cost is due to the civil engineering that is necessary to build or upgrade the customer access network. In Europe, for example, 60-80 percent of the overall investment costs into last mile access networks are due to civil works (Analysys Mason, 2008). In the US Verizon has to pay an estimated 4000$ per connected customer for its fiber product FiOS. Only 650$ of these costs are for labor and the equipment itself (Hansell, 2008). Such an investment structure causes periodical overprovisioning of bandwidth, which, however, is soon filled up again with new content. Just like John Maynard Keynes once said “supply creates it’s own demand”, IUs and CSPs always find and create new ways to use the existing bandwidth. This is the dilemma that is constantly brought forward by network operators in the NN debate.

Mobile telecommunications providers are even more in trouble. Radio waves can only deliver a limited amount of bandwidth per cell[2] and the network is therefore constrained on a regional basis depending on the distribution of cell towers. With this additional restriction it becomes even more challenging to guarantee a certain level of quality, because customers can move around with their access device(s) and consequently cause congestion at highly frequented

[1] 10^{18}bytes (If one assumes, that one megabyte of data can hold a big book of around 400 pages one exabyte of data translates into a trillion books (cf. Swanson and Gilder, 2008, p.4).)

[2] Depending on the technology that is used by the network provider. Newer technologies (e.g. Long Term Evolution (LTE)) provide higher throughput with the same limited frequency spectrum.

spots of the covered areas.[3] Another problem of mobile technologies is the unpredictable influence of environmental conditions on the quality of the connection. Whereas fixed access providers usually do not have to deal with the problem of interference, mobile carriers have to take into account the topology of the terrain, general weather conditions and the speed of customers traveling through the coverage area of a cell. This inherent uncertainty makes it very difficult to deliver reliable service and a high level of quality compared to fixed line access.

To cope with the increasing demand for data transportation Network Equipment Providers (NEPs) like 'Cisco', 'Alcatel Lucent' and 'Huawei' are constantly improving their products. Making use of new sophisticated multiplexing methods and additional colors in the spectrum of light, fiber optic networks currently cope with the increasing traffic in the Internet. As a result of this technological progress, costs per unit of bandwidth are decreasing.

One could ask why ISPs are simply not raising the prices for IUs up to a point where the revenue is in line with the resulting infrastructure costs, or introduce a pricing scheme that somehow corresponds to the data consumption of the customers. With respect to this argument ISPs argue that compared to 'flat rate pricing', so-called 'metered pricing' within some kind of two-part tariff is very unpopular and deviations from a flat rate regime back to metered pricing schemes are often very difficult to communicate to the customer base. In 2010 the big Canadian ISPs tried to revert to this pricing regime by imposing usage based billing on their wholesale products. All small ISPs relying on resale and wholesale products of the big Canadian ISPs would not be able to offer real flat rates anymore, since they themselves would be forced to buy usage based access products. With the whole country in jeopardy to loose unlimited Internet access, tremendous public protest arose and finally regulators decided to stop the larger telecommunications providers (Openmedia.ca, 2011). ISPs claim that the market for fixed Internet access is caught in some kind of 'flat rate trap'. While the prices

[3]Nevertheless mobile network providers currently profit from the general trend to mobile platforms and the increasing popularity of smart phones. In contrast to local access providers they still face an emerging market for Internet access.

are driven down by competition, customers would not accept a unilateral deviation back to metered tariffs. In addition to that, the demand per customer is higher under flat rate pricing and consequently causes higher operating costs for the ISP. In the context of Internet access pricing this phenomenon was also confirmed in the Internet Demand Experiment (INDEX) project[4]: "People were willing to pay a substantial premium to face un-metered pricing, but they also placed much larger demands on the system than when they were metered" (Varian, 2002, p.15).

This quote shows again the dilemma brought forward by the local access ISPs. On average consumers are willing to pay a premium for flat rate tariffs, but prices are driven down by competition and regulation, with the result that this additional willingness to pay cannot be internalized by the ISPs. In addition to that, flat rate customers consume more bandwidth and data, causing higher costs due to necessary additional investments in the network infrastructure.

Another problem of ISPs, closely related to flat rate pricing, is a very small fraction of customers ($\sim$1%) which is causing a significant share of the overall traffic. In mobile communications for instance, these so-called 'heavy users' or 'bandwidth hogs' accounted for 35 percent of the overall data traffic in 2010 (Cisco, 2012). It is obvious that flat rate pricing results in some consumers paying too much compared to an alternative metered tariff, whereas heavy users are subsidized by this group. How ISPs currently handle heavy users in their networks will be discussed in detail in Section 2.1.1.

In summary, ISPs claim that their investments in the network are hardly counter-balanced by new revenues from IUs. In reverse, CSPs benefit from the increased bandwidth of the access networks, which enables them to offer even more bandwidth demanding services to their customers. Not surprisingly ISPs argue that their current business model cannot be sustainable anymore in the near future and a shift towards new pricing and network management

[4]Within the INDEX project a large field experiment with Internet users living in university dorms was conducted. Their demand for bandwidth and data was measured in reaction to different pricing schemes for Internet access.

regimes should not be blocked by regulation to ensure fast and reliable Internet access for all customers.

However, the open question remains if the overall net effect on ISPs revenues will indeed be negative and if their investments into the network infrastructure under a neutral network regime will be sufficient or not.

1.4. Stakeholder and Revenue Streams in the Internet

1.4.1. Internet Users

Revenues in the Internet ecosystem are primarily generated by the IUs. In many countries customers predominantly pay flat fees for their access to the Internet. Very often the access service is differentiated solely by the maximum possible throughput of the physical connection.[5] If no flat fees are offered, customers usually pay fees for being connected to the network and additional usage fees according to the transferred amount of data or according to the time they are being connected to the network.

In addition to that IUs generate revenues on the CSP side of the market through advertisement impressions or direct payments. Even under the current trend to deliver content and services via hardware tailored applications (e. g. , Apple's AppStore), rather than standard based web sites or formats, the biggest share of revenues in the Internet economy today is still created by advertisements (Dou, 2004). The Internet business is subsequently often called a 'competition for eyeballs'.

[5]Customers can, for instance, decide between Digital Subscriber Line (DSL) service with six megabit per second or service with 16 megabit per second and pay a premium for the flat rate with higher bandwidth.

Without NN regulation it would be possible to charge IUs for preferential access to the network or the allowance to use certain services.[6] All IUs have different preferences and therefore differing usage patterns and requirements. A business user for instance could be charged a premium to get reliable and fast access to the network even in times of congestion. The additional black arrows in Figure 1.1 indicate, that the respective revenue streams (IU to ISP) could change under a NNN regime. This specific scenario will be further analyzed in Section 2.3.2.

1.4.2. Internet Service Providers

From an economic point of view ISPs are platform operators that connect the supplier of content and services with consumers through their access networks. In the NN debate the terms ‘core’ and ‘edge’ of the Internet are frequently used. As Figure 1.1 shows, the term core describes all hardware and physical connections that are used to transport data from CSPs to IUs. The core falls into three subcategories.

Access network. The physical connection to each household.

Backhaul. This network segment aggregates the traffic from all connected households of a single ISP.

Backbone. This is the highest level of traffic aggregation. It describes the interconnection of different network segments owned by different entities.

The local access networks are operated by so-called eyeball ISPs. Interconnection between the backbone and eyeball ISPs is warranted by a set of mutual agreements that are either based on bill-and-keep arrangements (peering) or volume-based tariffs (transit). In case of transit, the eyeball ISP has to pay the backbone ISP, and not the other way around. The edge of

[6] T-Mobile Germany offers a ‘Voice over IP option’, that includes nothing more than the allowance of the network operator to use services of this kind.

the network usually describes the periphery of companies and users that can connect to the network platform and who are able to innovate on the basis of this access technology (i. e., the network platform). In the Internet economy, access to the network is never free of charge. All entities pay for transportation of data through the network.

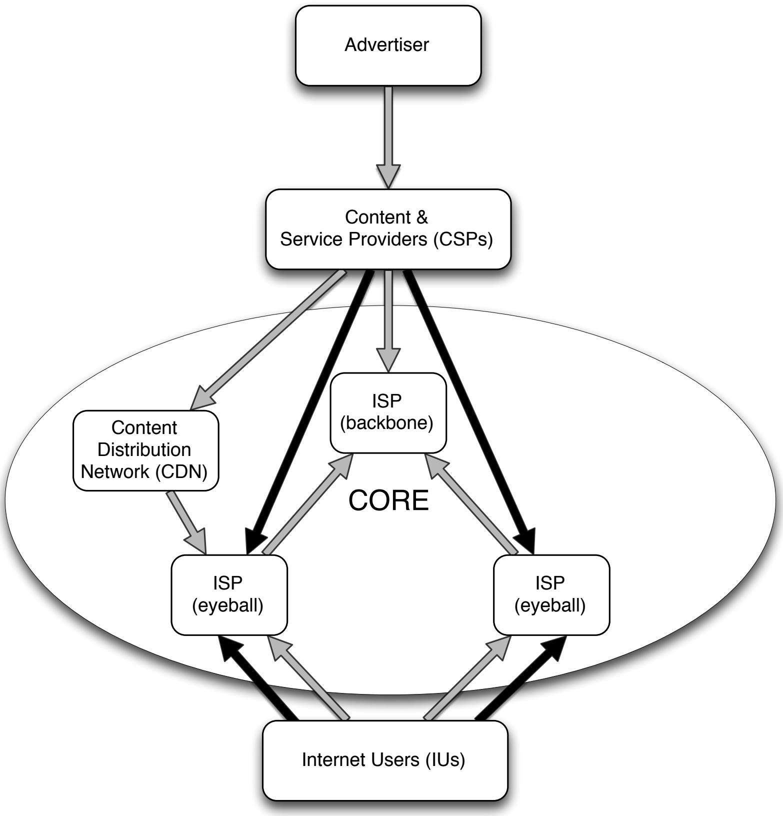

Figure 1.1.: Revenue streams in the Internet

1.4.3. Content and Service Providers

As mentioned before, CSPs generate revenues predominantly by advertisements. The advertisement revenues increase with the number of users visiting the CSP. The level of advertisement revenue depends on the click-through rate of customers (i. e., the proportion of arriving customers clicking on an advertising banner) and the revenue per advertisement impression

(i. e. , the revenue per arriving consumer who potentially looks at the banner). The revenues of CSPs are tightly interweaved with the revenues of rights owners. It is of course possible that rights owners are also active in the role of a CSP itself, but usually they make money with the royalties of the content they produce (e. g. , TV shows, movies, music).

CSPs are paying for access to the network, too, but usually not to the terminating ISP who delivers the content to the IUs visiting the respective CSP. This is very important to understand. In the current state of the Internet ecosystem no additional fees are imposed by eyeball ISPs to any CSP. Therefore, from the point of view of a last mile network operator CSPs are currently 'free riding' on their infrastructure

Without NN regulation this condition could change in the near future. It would be possible for ISPs to charge CSP's a positive price just for access to their networks. This is indicated by the additional black arrows in Figure 1.1. This would be a whole new source of revenues for eyeball ISPs, without changing anything on the underlying network architecture, but merely requiring an additional billing system for CSPs. This termination fee scenario will be further discussed in Section 2.2.

Obviously revenues of CSPs are somehow related to the value and the performance of the content or the services they offer to IUs. If consumers experience a bad service quality due to congestion in the network, they will probably not visit this CSP again, because they attribute the bad quality directly to the CSP. For this reason CSPs are eager to improve their Quality of Experience (QoE) for the customer. It has been already explained before that the status quo of data transportation in the Internet follows for the most part the BE principle. This means that all data packets are usually transported with the same priority through the network and there is consequently only one class of service. BE implies that very demanding services suffer more in times of congestion than other services. The current Internet has no intelligent network mechanisms to circumvent this and to deliver higher QoE. This is the downside of the BE principle. Without NN regulation it would be possible for eyeball ISPs to offer multiple service classes and grant preferential access to their customer base via QoS. The additional black arrows in Figure 1.1 indicate that this could also be a whole new source of revenues for ISPs (instead of charging only for bare access to the network itself without deviating from the BE principle). This NNN scenario of CSP tiering will be further discussed in Section 2.3.1.

As Figure 1.2 shows the QoE of a CSP is influenced by three major dimensions:

- Content or service requirements w. r. t. to QoS
- End to end connection quality
- User preferences and expectations about service experience

To deliver content and services with a higher QoE, CSPs currently commission Content Distribution Networks (CDNs) and rely on their caching and route optimization services. QoS tiering (i. e. , another shape of NNN) is one additional possible way to account for these individual service demands and deliver certain data packets with priority.

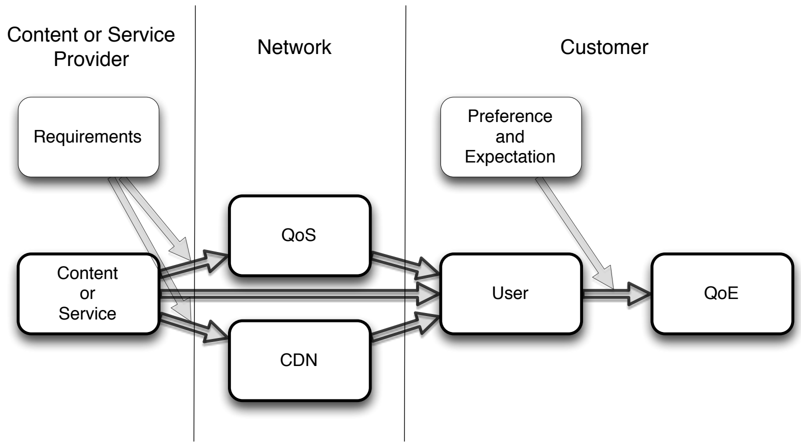

Figure 1.2.: Quality of Experience (inspired by Kilkki, 2008)

1.4.4. Content Distribution Networks

In the current BE Internet, where prioritization of data is not used, companies like 'Akamai', 'Level3' and 'Limelight' earn money by selling caching services and optimized routes through the Internet to CSPs. Their service can be considered as a tool to deliver content and services with higher quality to the customers, as can be seen in Figure 1.2. This business model makes it necessary to build additional infrastructure to bypass congested routes on the public Internet by caching frequently downloaded content closer to the requesting IUs. CDNs are an essential and very important part of todays Internet infrastructure. Akamai for instance claims that it delivers between 15 and 30 percent of the overall Internet traffic in the world. Figure 1.3 shows a snap-shot of the live statistics of the Akamai CDN. At the time of the snap-shot, Akamai delivered about two million streams to IUs worldwide and experienced about 18 million Hyper Text Transfer Protocol (HTTP) hits per second (each hit followed by delivery of non-streaming related data). In Section 5.1 the role of CDNs in the NN debate will be discussed in more detail.

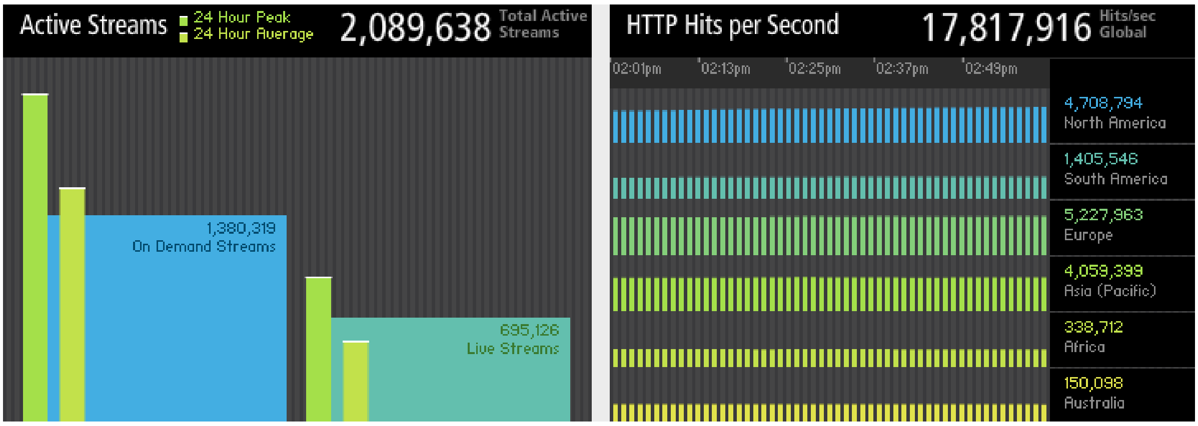

Figure 1.3.: Akamai live statistics (02/12/2012)

Chapter 2.

A Non Net Neutrality Framework

"People who demand neutrality in any situation are usually not neutral but in favor of the status quo."

Max Forrester Eastman

THE previous Chapter explained in detail what can be subsumed under the term NN and where in the history of the Internet the NN debate is rooted. It has been analyzed where ISPs see possible new sources of revenue in the commercial Internet ecosystem in the future. It is obvious that access to customers is key in the Internet economy. Big eyeball ISPs are gatekeepers of their installed base and consequently to the mass of potential customers of CSPs. In this chapter it will be discussed how ISPs handle data transportation today and how they could exert more control over this data. Along these lines the relevant possible NNN scenarios are evaluated.

In the long lasting debate many different scenarios of NNN have been studied and potential dangers to the status quo of the Internet have been analyzed. Currently there seems to be consensus about the most critical points that have to be addressed, but there are still many discrepancies about the most likely scenario of NNN, which obviously crucially influences the implications for possible NN regulation.

There are only three notable papers which structure the NN debate and the related literature. Schuett (2010) provided the first review that focuses on the presentation of the theoretical economic literature. In particular, Schuett distinguishes between NN as a *zero pricing rule* and as a *non-discrimination rule*. This categorization inspired parts of the following framework. Faulhaber (2011) gives a more general introduction to the debate and also discusses the relevant literature that has emerged in the post Schuett (2010)-era. He also analyzes the new Federal Communications Commission (FCC) rules and draws the clear cut conclusion that no evidence or academic result would justify ex-ante NN regulation. Bauer (2007) analyzes first and foremost regulatory measures and identifies three possible governance structures that could deal with potential NNN threats: Relying fully on anti-trust law, non-discrimination rules and full NN regulation. His policy implications inspired parts of the structured policy decision guideline that is presented in Section 2.4.

Figure 2.1 presents a new framework to structure the NN debate that combines and extends previous approaches. In particular, two general dimensions are identified that form all possible NNN scenarios. In this vein, not only economic papers, but also literature from the law and engineering domain can be categorized. Especially in the engineering domain network management and prioritization mechanisms had been studied long before the NN debate emerged.

		Pricing regime	
		One-sided	**Two-sided**
Network regime	**Quality of Service**	Internet User tiering (IUs choose priority class.)	Content and Service Provider tiering (CSPs or CSPs&IUs choose priority class.)
	Managed network	Status quo (Best Effort network with traffic engineering and/or managed services.)	Termination fee (Additional fee for CSPs to terminate traffic at eyeball ISP.)
	Capacity only	Strict Net Neutrality (No discrimination based on source, destination or content.)	

Figure 2.1.: Non Net Neutrality framework

2.1. Neutrality in Internet Access?

2.1.1. The Status Quo

Contrary to the believe of some proponents of NN, there are already management and prioritization mechanisms for data packets implemented in the network infrastructure of eyeball ISPs today. The current generation of hardware is already able to support QoS, and new wireless standards like LTE, the next generation of mobile data communication services, have similar mechanisms already incorporated into the specification itself. One best practice of eyeball ISPs nowadays is called 'traffic management' or 'traffic engineering'. This procedure describes the intelligent management of traffic flows and aims mainly at the reduction of infrastructure costs and to ensure service quality for the better part of the customer base. Traffic management practices can be used to either throttle specific users or protocols[1] (e. g. , Peer-to-Peer (P2P)), or to throttle in peak load situations (Dischinger et al., 2010). The following list from Dischinger et al., p.2 shows how differentiation can be performed "[...] by examining one of the following:

[1]In technical terms this is referred to as 'flow based differentiation'. A 'traffic flow' describes the whole of data belonging to a specific type of application.

(a) *The Internet Protocol (IP) header:*

The source or destination addresses can determine how an ISP treats a flow. [...]

(b) *The transport protocol header:*

ISPs can use port numbers or other transport protocol identifiers to determine a flow's treatment. For example, P2P traffic is sometimes identified based on its port numbers.

(c) *The packet payload:*[2].

ISPs can use Deep Packet Inspection (DPI) to identify which application generated a packet. For example, ISPs look for P2P protocol messages in packet payload to rate-limit the traffic of P2P applications, such as BitTorrent."

Based on these techniques the ISP can decide how to handle the identified data packets. Without going into too much technical details, one can distinguish between two possible outcomes that can be achieved by technically handling some data packets differently: Packets of certain applications, services or content are not delivered to the requesting customer (*blocking*). The experience while using or consuming certain applications, services or content is impaired (*degradation*) (Dischinger et al., 2010). In an interview Georg Merdian, director of the infrastructure regulation devision of Kabel Deutschland, Germany's largest cable company, said that the cable network currently is sufficient for the data traffic generated by the customers, but "We anticipate we will soon have to use some kind of management techniques"(O'Brien, 2010).

[2]A data packet consists of control information and user data. The user data of a packet is also called the 'payload'.

The potential dangers of mechanisms like this and the resulting fears among NN proponents are obvious. As described in Section 1.4.1, heavy users account for a large proportion of traffic in the networks. ISPs could use this technology to artificially degrade the traffic of such users to save costs, even if they are paying for an unlimited flat rate plan.
Likewise ISPs could argue that these procedures are necessary to ensure a desired level of QoE to the better part of customers, while the true reason may be anticompetitive behavior such as foreclosure of other revenue sources, e. g. , preventing Voice over IP (VoIP) services to deliver the same service quality as the own telephony service.
In the US two prominent examples for the abuse of traffic management are 'Madison River Communications'(2005) and 'Comcast'(2008). Madison River Communications artificially degraded other VoIP providers to make it's own service more attractive. Comcast on the other hand degraded P2P traffic to reduce congestion in it's network during peak times. While the first case ended with Madison River Communications paying a relatively low fine ($ 15,000), Comcast challenged the authority of the FCC and won the legal dispute (Faulhaber, 2011). This shows that consumer rights and anti-trust disputes are not the legal core competence of network industry regulators, but that NN would extend the scope of regulation to these grounds.

A similar argument may hold true for so-called managed services (e. g. , commercial Internet Protocol Television (IPTV) offers). Managed services are delivered over the same IP based connection to the customer as other Internet data. To guarantee service quality, ISPs give IPTV packets preferential treatment in their networks, but very often also operate their own hardware to store and distribute the content. It is not surprising that almost all of these offers are bundled with the network access products of the respective ISPs.[3] ISPs could have an incentive to foreclose rival service providers from essential parts of their network architecture to ensure premium service quality exclusively for their own products.
Another drawback of such preferential treatment is the negative impact on the transportation of other Internet data. During a bottleneck situation in the backhaul, the ISP ensures firstly that

[3]ISPs came up with these products against the competition of cable operators, which offer TV services as their (formerly) primary business model.

all premium services are running smoothly, while customers requesting data from the Internet may experience a degradation in service quality.

Another category of fears is related to free speech and freedom of information. Especially consumer rights groups and independent journalists are afraid of ISPs being able to decide which content and opinions are available to their customers. ISPs could block access to political controversial (but legal) information, or shut down websites of unwanted organizations (e. g., the websites of labor associations to prevent an assembly of workers). According to strict definitions of NN like Definition 1.1 all forms of network management described above cannot be considered as neutral anymore. However, the status quo of the Internet can in large part be described as a de facto managed BE network regime.

Many law scholars focus on the potential dangers under the status-quo. Especially Wu (2003) discusses in detail the differences between necessary network management and harmful degradation, and he proposes a NN law called 'Forbidding Broadband Discrimination'. His formulation is very similar to §63 and §68 of the FCC ruling (see Section 2.4), because he emphasizes the right to reasonably use the Internet connection, but also accounts for the necessity to ensure the quality of the broadband service for the better part of the customers. Christopher Yoo (2005), also a law professor, can be seen as his dogmatic counterpart. Whereas Wu is concerned with the gatekeeper position and anti competitive behavior of ISPs, Yoo highlights the efficiency gains of QoS and believes in higher infrastructure competition under differentiated services. In his view differentiation facilitates that more ISPs survive and therefore more alternatives in local Internet access exist. Both authors also differ in their opinion about innovation in the Internet ecosystem. While Wu believes that innovation at the edge of the network is more important and cannot be compensated by the innovation at the core and new local access products, Yoo thinks that the natural monopoly theory has led to the false premise that competition in the service layer is more important than competition in the infrastructure layer of the Internet ecosystem. In his eyes NN is a matter between the big CSPs and big ISPs.

Whereas Wu and Yoo focus on the aspect of traffic management, QoS and price discrimination, van Schewick (2007) analyzes the incentive of ISPs to discriminate against unaffiliated CSPs of complementary products. She concludes that NN regulation is necessary to protect independent producers, but this comes not without costs. Van Schewick sees a direct trade-off between innovation at the network level (core) and innovation at the application level (edge). In her view applications are the main driver of the Internet economy and therefore innovation at the edge of the network is more important. Consequently NN regulation is needed to foster this innovation. She comes to this conclusion because she assumes that the reduction of innovation incentives of a high number of potential innovators can not be compensated by the higher innovation of a few network providers. Additionally, she claims that centralized innovators like ISPs cannot successfully replicate with their own services the innovative potential of a large number of independent developers. Although the paper follows a convincing line of reasoning, this appears to be more an assumption than a conclusion.

The question remains whether potential anticompetitive degradation of competing services or nontransparent throttling of users is already a unique NN problem. Most of the actions described above are already illegal under the status quo. Constitutional law stipulates the right of free speech. Likewise antitrust law identifies artificial degradation of service to rivals as illegal.
The lack of transparency and the resulting lack of assessability of the possible negative effects of such practices in reality is one of the biggest problems of the status quo. Likewise, transparency could also be one potential remedy to this situation.

One approach in empowering users to detect whether their ISP is differentiating between types of services of specific applications is the ‘Glasnost’ project (Dischinger et al., 2010). The research project at the Max Planck Institute for Software Systems aimed at creating an online tool that enables IUs to check whether their ISP is actually interfering with their data transmissions. They found that on average at least 10% of users experienced degradation in P2P traffic and that contrary to the arguments of the ISPs this happened throughout the whole day and not only at peak times. It is even more surprising, as Table 2.1 reveals, that Kabel Deutschland was using traffic management procedures on a large scale even before the interview with Georg Merdian took place.

Table 2.1.: DPI of German ISPs based on Glasnost data in Q4/2009

Operators Name	**Number of valid tests**	**Percent of tests showing DPI**
Kabel Deutschland	250	39%
Deutsche Telekom	205	3%
Vodafone Germany	116	4%
HanseNet Telekommunikation	112	7%
Telefonica O2 Germany	50	2%
Kabel BW	27	7%
Unitymedia	26	4%
NetCologne	18	11%
Versatel Communications	18	6%

(Mueller et al., 2009)

Beside these bottom-up tools, that do not rely on the fact that the ISP is handing out more information about the network management practices, one could also think about forcing ISPs to report certain quality related measures to the connected customers (top-down) directly. The usefulness of transparency in the context of NN has already been discussed before by Faulhaber (2010) and Sluijs, Schuett, and Henze (2010). Faulhaber claims that information has to be easily accessible and understandable to be helpful in the NN context. He draws a comparison to nutrition information on groceries. Firstly, there has to be information available on the product, otherwise consumers incur unnecessary search costs. If the information is accessible, but complex and difficult to understand, the information does not help consumers to make a more informed decision. In addition to that the information should be verifiable. In contrast to this qualitative approach, Sluijs et al. (2010) use the methodology of an economic laboratory experiment. They simulate a market with two ISPs and a potential customer base with human participants, varying the information about the delivered service quality available to the customers. Their most important result is, that already a fraction of informed users can help the whole market to achieve a welfare superior outcome. This means that as long as enough consumers are able to obtain the correct information about the service quality, welfare is increased.

The aspect of transparency is closely related to the aspect of competition in local access provision. In case customers indeed find out that their access provider throttles bandwidth in peak times, or degrades certain services, they can only switch the provider in the face of competition. Otherwise their option is to stay with their respective ISP, or to live without any Internet access. This scenario perhaps sounds strange and unrealistic, but many rural areas (especially in the US) are only served by a maximum of two ISPs. Usually one of them is the local telecommunications provider, while the other is the local cable company. Opponents of NN regulation argue that abuse of network management as well as other deviations from the status-quo are unproblematic in the face of competition.
Wu (2007) analyzes the US mobile phone market with respect to NN and finds many examples of non-neutral behavior (e. g. , crippled products and degradation) in an, as is often claimed, highly competitive environment. He explains the interplay of competition and transparency: "To say that competition can then be a reason not to examine industry practices and mandate as much disclosure as possible is exactly backward. For it is such information that is necessary to make competition work in the first place." (Wu, 2007, p.423)
However, competition may not always be beneficial under NNN cenarios. Kocsis and Bijl (2007) argue that termination fees and exclusive deals with CSPs can lead to more horizontal differentiation of ISPs and consequently to higher mark-ups. This conversely relaxes competition between the ISPs in the market.

2.1.2. Strict Net Neutrality Scenario

Strict NN (as postulated by Definition 1.1) would imply a deviation from the current status quo of the Internet towards a network regime where any network management practice would be forbidden. This regulated technical disarming could lead to congestion problems, the cementation of costly overprovisioning and reduced revenues, because the business with managed services like IPTV would be in danger.

Another drawback is that strict NN prohibits to offer limited Internet access for a lower price. This could mean anything from access to the Internet with lower priority in the backhaul of the ISP, to unhampered access to only a subset of content or services for a lower price. This line of argumentation is also acknowledged by the vice president of the European commission Neelie Kroes who said that: "[...]requiring operators to provide only 'full internet' could kill innovative new offers [...] Even worse, it could mean higher prices for those consumers with more limited needs who were ready to accept a cheaper, limited package" (Meyer, 2011). For these reasons it is more than unlikely that ex-ante regulation of strict NN can be considered useful and welfare increasing. Nevertheless, countries like Chile walked down this road in the year 2010. Chile was the first country in the world that intended to codify NN in law. According to this law even network management should be forbidden and would fall into the category of unlawful interfering with Internet traffic. The regulation of NN in general and the example of Chile will be further discussed in Section 2.4.

The following two sections will extend the picture of NNN along two additional dimensions. Beyond the debate whether the status quo of the Internet and the behavior of ISPs can be considered as neutral or not, two general directions of deviating from the status quo are discussed in the literature, forming all relevant NNN scenarios.

2.2. Two-Sided Pricing

According to the two sided market literature (e. g., Rochet and Tirole, 2006; Armstrong, 2006) each eyeball ISP can be considered as a platform provider in a two sided market. Following the logic of this economic concept, each ISP is some sort of match maker between two groups of customers located on the opposite sides of the marketplace. In a two-sided market, it is possible to charge customers on both sides of the market for bare access to the platform or (and) for usage of the platform's match making process. Prominent examples of two-sided markets are partner matching services (women/men), credit cards (cardholders/shops accepting the card), gaming consoles (console owners/game publishers) and online auction services (buyer/sellers).

One important aspect of two-sided markets are network effects. Network effects in their direct form describe the fact that participants of a network have a positive valuation for the number of other subscribers. By contrast, two-sided markets exhibit a general form of indirect network effects. Hagiu (2006) coins the term 'bilateral indirect network effects', while Rochet and Tirole (2006) call them 'cross side network effects'. However one prefers to label the effect at hand, this means that each side of the market has a positive valuation of the network size on the other side of the market (membership externalities) and additionally also for the usage of the platform from the other side (usage externalities).[4]

In terms of the Internet, one has to recall the revenue streams described in Section 1.4. IUs value to be connected to the Internet and they also value variety of content and services as well as their usage. The inverse holds true for CSPs and their potential advertising revenues generated by users looking at or clicking on banners. However, in addition to the 'classical' view of network effects in two-sided markets, IUs may also prefer that other users are connected to the platform, because they can communicate with them (e. g. , via a messaging service).[5]

[4] According to this definition, indirect network effects are only a special case of a two-sided market with membership externalities, but without any usage externalities (Rochet and Tirole, 2006).

[5] It has to be noted, that the same argument holds true for some classic examples of two-sided markets as well. If one considers gaming consoles, one typical example of a two-sided market, one can easily comprehend the valuation of the two sides for each other. Gamers value the variety of games and consequently the commitment of developers and publishers for the console. Developers and publishers like a large installed base and the presumably higher sales rate of titles for a successful console. However, with the connection to the Internet and the rise of online multiplayer games, users of the console itself value the existence of other players on the same console type positively. In reality two-sided markets like that exert a combination of direct and cross side network effects.

2.2.1. Termination Fee Scenario

According to this economic concept, one possible NNN scenario could be that eyeball ISPs consider themselves as platform operators connecting CSPs with the IUs via their network. From a two-sided market perspective one side of the eyeball ISP platform currently enjoys a zero price. Under this new regime, ISPs could use their market power over the last mile customers to charge the CSPs additional fees to connect to the installed base. However, it is important to note that the data transportation method under this scenario does not change, meaning that all data packets are transported following the BE rule.[6] Even among the papers assuming this scenario of NNN there is no consensus on how CSPs are charged for access to the network of the eyeball ISP. Theoretically, any possible combination of membership and usage fees is possible.

Besides the underlying *pricing structure* for network access, proponents of NN are even more concerned about the aspect of *price discrimination*. In case termination fees would be allowed on an individual CSP level, this would be first order price discrimination. With very few exceptions, there is consensus that price discrimination techniques, which are not beneficial to the IU side of the market, should not be allowed on the CSP side of the market either. Therefore, even many opponents of NN vote for some kind of 'non discriminatory surcharge' rule in pricing.

However, there are different forms of non discriminatory access possible. For instance, ISPs could charge all services using the same protocol (e.g. Session Internet Protocol (SIP)) or that are in the same group of services (e.g. VoIP) the same fees for data transportation, but services in another class (e.g. IPTV) a different price. Because services in different classes are usually not in direct competition with each other and the quality expectations of consumers may differ between the two types of services[7], this price discrimination procedure could be considered

[6] Or at least to the same conditions as described in Section 2.1.1

[7] IU usually expect more reliability and quality with real time communication services than with streaming media, where buffering of data can already account for temporary quality reduction due to network congestion.

as non discriminatory. Interestingly, following NN Definition 1.4 such behavior would be considered as in line with NN by Google. This shows again that huge differences between the views of the different stakeholders in the debate exist. However, the most undisputed pricing scheme would be an equal fee for all CSPs that want to connect to the network of an eyeball ISP.

First, it is important to understand that the termination fee model, as well as all other NNN models to be discussed, will generally pose the same concerns as the current status quo. In other words, if ISPs have an incentive to block or degrade costly traffic flows or heavy users under a managed network regime, why should this incentive not prevail under a two-sided pricing or QoS tiering scenario? The same logic applies to concerns about freedom of speech. If ISPs would indeed want to block e. g. , the websites of labor associations to prevent an assembly of workers, why should they not do so under another network regime?
If they indeed pursue such goals, the network regime is not of importance to this matter, as long as the network technology offers the possibility to differentiate data packets. This holds true for all scenarios except for strict NN. Another aspect mentioned before is possible anticompetitive behavior against rival content or services. If an ISP has affiliated content and he uses his control over the network to artificially degrade or block rival content for his customer base, he could do so likewise in all scenarios except for strict NN.

Nevertheless, two-sided pricing brings the additional potential pitfall of reduced content innovation on the table. Proponents of NN argue that already a positive price may reduce content variety, because some CSP could not afford the money to buy access to some ISPs.

Economides and Tag (2012) were the first to present an economic model in the context of the NN debate. They consider the termination fee model where CSPs pay a lump-sum to connect to the IUs. The essential assumption in the model is that CSPs value an additional IU more than IUs an additional CSP. The first version of the paper was available in 2007 and many things changed in the process of refining the model. The main finding in the published version of the paper is that IUs and ISPs are better off with NNN. Regulators, who are often most concerned with consumer welfare, should therefore be considerate before imposing mandatory NN. This result is a direct consequence of the above assumption and fully in line with the extant two-sided market literature: The more consumers can be attracted, the more profit can be generated with additional fees on the CSP side of the market. Consequently, consumers enjoy a low subscription price under NNN and ISPs are allowed to extract additional revenues from the CSPs. Under monopoly NN is only welfare enhancing if the differentiation between the consumers is relatively high. In other words, this would mean that IUs have a very strong brand preference for a particular ISP compared to their valuation of access to and content in the network. This is a very questionable case in the context of a homogeneous product like Internet access. Therefore, in their model only CSPs would profit undoubtedly from NN regulation. Although the results of the preliminary version of the paper supported the need for NN regulation, the results of the published version are therefore rather tipped in favor of NNN. The authors themselves conclude that the welfare results are ambiguous. Nevertheless, the published version as a whole is written in a very NN orientated manner.

Njoroge, Ozdaglar, Stier-Moses, and Weintraub (2010) follow a similar approach, but they add the platform investment decision and interconnection between two ISPs to the picture. Both platforms charge flat access charges to CSPs and IUs. Under NN the platforms differentiate maximally resulting in a high and low quality platform. They show that welfare in their model is generally higher in the NNN regime, because the NNN regime leads to higher infrastructure investments by the low quality ISP. In their model the same argument holds true for consumer surplus and CSPs revenues. CSPs revenues increase through higher advertising revenues, overcompensating for the higher price for access. Although the welfare results are unambiguous, it is interesting that the high quality ISP prefers the NN regime. This is due to the fact that under the NNN regime the low quality ISP can catch up through additional investments in the network infrastructure, resulting in fiercer competition and lower revenues.

Musacchio, Schwartz, and Walrand (2009) incorporate also investment costs into their model, but mainly add to the debate by exploring the effect of multiple ISPs charging for access.

ISPs in their model are not in direct competition with each other, but rather forming coexisting spatial monopolies. They are interested in the price charging behavior in this situation. The authors show that two-sided pricing is preferable if the ratio of advertising rates to price sensitivity is extreme. However, otherwise a situation similar to the tragedy of the commons may arise in equilibrium. ISPs tend to ignore their own negative effect of overcharging on the investments of the CSPs and consequently on the revenue of all other ISPs. This negative effect becomes more prominent as the number of ISPs increases; resulting in NN becoming more attractive.

Hermalin and Katz (2007), mainly interested in studying product line restrictions, apply their model to the NN debate. They analyze a monopolistic ISP offering a menu of qualities, because he can not observe the valuation for quality of each CSP. In a special case of their model they look into the enforcement of a zero-price rule in which the ISP would only produce one quality. This quality would be even lower compared to a regime where only one quality would be enforced, but positive prices to CSPs are allowed.

The first article by Wu (2003) focused exclusively on the status-quo of the Internet, whereas Lee and Wu (2009) discuss two-sided pricing and argue in favor of a zero-price rule for CSPs. Firstly, they highlight the important fact that all IU are potential future CSPs and that a zero-price rule ensures cheap market entry. Also, no one has to ask for permission to reach the installed base of ISPs. In addition to that they make the point that eyeball ISPs are actually already compensated for traffic by peering and transit agreements. Since these contracts are negotiable and voluntary, there is no reason why higher infrastructure costs could not be supported by more favorable agreements. A zero-price rule would also eliminate transaction costs, because two-sided pricing makes new billing and accounting infrastructure for CSPs necessary, and because it introduces a new form of transaction costs as well. The most striking argument in the paper deals with the potential fragmentation of the Internet ecosystem. If two-sided pricing would result in CSPs deciding whether they want to connect with each individual ISP for a given price, than it seems inevitable that IUs will have only access to a fraction of the CSPs they have today, depending on the ISP and his agreements with CSPs.

Yoo (2005) argues that the 'burden of proof' that NNN is indeed worse than NN is on the side of those who want to regulate. He therefore calls for an 'wait and see' approach. Yoo doubts that the assumption put forth by proponents of NN, that bandwidth increases faster than demand, is correct. In his opinion overprovisioning[8] is more likely to be inferior compared to managing or diversifying the network.

Many of the papers mentioned before try to argument against or in favor of NN regulation. The question remains what policymakers and regulators could do about the potential dangers under a termination fee scenario, besides enforcing a NN regime and possibly risk reduced infrastructure investments in the future (see Section 2.4).

As already discussed, one important piece of the puzzle is price discrimination. As long as ISPs would be allowed to e. g. , auction of access (Choi and Kim, 2010), financially strong CSPs would have an indisputable advantage. Therefore, a *non discriminatory surcharge* may be one regulatory tool to ensure a level playing field for all CSPs. Remember that different forms of non discrimination are imaginable.

However, this may not fully solve the problem, because the main argument is not invalidated by this tool, namely the possible exclusion of young start-up companies with a constrained budget. Nevertheless these companies have to raise money for all other aspects of their business as well (e. g. , hardware, personnel, etc.). One viable way would be to grant promising companies access to the network (or likewise under a QoS regime the priority lane, see also Section 2.3.1) free of charge. Like these businesses have to pitch for money to start their business, they could pitch their idea to get access to the network (the priority lane). This argument is often brought forward by ISPs, however, it is exactly this kind of ex-ante selection of ideas that proponents of NN perceive as dangerous to the open spirit of the Internet.

[8] 'Overprovisioning' describes the approach of network owners to plan network resources with enough spare capacity to guarantee the highest service quality even under peak loads.

Another option would be some kind of non-discriminatory revenue sharing. Start-up companies could commit themselves to pay a fixed share (for a limited time) of all potential revenues to the ISP. This would not hinder access to the network for anyone, but allow the ISP to generate additional revenues in the case the business idea becomes successful. This construct would allow all possible ideas to be challenged by the market and fees would only be imposed in case the innovation becomes successful.

2.3. Quality of Service

QoS as a general concept in communications networks allows to transport telephony or data traffic with special requirements. In contrast to traffic management policies, which usually deal with ISP internal resource optimization, QoS tiering describes a network regime based on the assignment of different priorities to data packets in general. Priority can be assigned on the CSP level, but also on the IU level.

Currently the infrastructure in the Internet (e. g. , router) works on a BE basis, but the Internet Protocol Version 4 (IPv4) already contains a Type of Service (TOS) field in its header by which routers could prioritize packets in their queues and thereby establish QoS. However, a general agreement on how to handle data with different TOS entries was never reached and thus the TOS field was not used accordingly. In telecommunications engineering, research on new protocols and mechanisms to enable QoS in the Internet has been done long before the NN debate evolved. In context to the NN debate, the DiffServ [RFC 2474] architecture is often discussed because it allows for a relatively simple class based differentiation of traffic. Also the current Internet Protocol Version 6 (IPv6), which was standardized in 1998, contains header information on the traffic class as well as a flow label, which facilitates QoS for real-time applications. In addition, data packets can even be differentiated solely based on what type of data they are carrying, without the need for an explicit marking in the protocol header.

This is possible by means of so-called DPI. All of these features are currently deployed in the Internet and could be used to shift the Internet access business to a new QoS tiering regime.

Such pay for priority arrangements seem less obtrusive than termination fees. If the non priority lane is still offered free of charge, all CSPs can at least connect with each ISP. However, it should be clear that given a fixed amount of bandwidth, speeding up some CSPs will inevitably lead to a slowing down of those CSPs that do not pay the priority fee. In a $M/M/1$ queuing model this translates to introducing an additional queue which handles the request of the CSPs in the priority class and which is processed ahead of the queue for the BE class (Kleinrock, 1976). However, in each class the queue is cleared on a first-come first-served basis. Assume that μ denotes the available capacity in the network and λ the traffic. The average waiting time in a neutral network is consequently denoted by $w_{NN} = \frac{1}{\mu-\lambda}$, which can be interpreted as the level of congestion in the network. However, now consider a tiered system in which the data packets of some CSPs are transported with priority, i. e., these packets are always enqueued ahead of the BE packets. If x denotes the share of all CSPs buying priority access, the average waiting time in the priority class is defined by w_P, and in the BE class by w_{BE}.

$$w_P = \frac{1}{\mu - x\lambda}, \qquad w_{BE} = \frac{\mu}{\mu - \lambda} w_P$$

It is easy to see that relation $w_P < w_{NN} < w_{BE}$ is always fulfilled, assuming an equal transmission capacity in both regimes and that not all CSPs are buying priority access. Otherwise, when all CSPs are in the priority class, the model trivially collapses to $w_P = w_{NN}$. This is an important feature of this queuing model, because it shows formally that serving some CSPs or IUs with priority will unambiguously lead to a degradation of service quality for the remaining network participants in the BE class. Chapter 3 will make extensive use of this approach to model congestion in the Internet ecosystem. For now consider only this basic trade-off, which is driving many results in the economic literature about QoS tiering on either side of the market.

2.3.1. Content and Service Provider Tiering Scenario

Many opponents of NN refer to the argument that the current BE Internet cannot be considered as 'neutral' since different types of data and applications have different requirements for network quality. Consider for instance an application like Skype.[9] In a situation where the network is congested, the resulting delay of data packages of the Skype service has a highly detrimental effect on the usability of the service itself. In comparison, the detrimental effect on an email service is negligible. According to this logic, highly congestion sensitive services are put at a disadvantage in a neutral BE network, because the quality of experience is lower in a bottleneck situation compared to a less demanding service. The introduction of QoS could mitigate the experience of such inherent differences in data transportation requirements of services. QoS allows the ISP to differentiate data packets based on different criteria and also to charge prices accordingly. Following this logic, highly congestion sensitive services could be better off if they were allowed to pay for priority, whereas the detrimental effect of a worsening in congestion to the remaining BE class could be less severe, because only congestion insensitive CSPs remain in this class. Again any combination of membership and usage fees for priority transmission would be possible, but as well for access to the remaining BE class. Also the forms of price discrimination correspond to the ones presented in Section 2.2.1.

However, QoS tiering bears one additional risk compared to the termination fee scenario. Through the additional source of revenues (selling priority access) ISPs may have the incentive to artificially degrade the BE quality to force more CSPs into the costly priority lane. This is also called the 'dirty road fallacy' (Sidak and Teece, 2010). This problem may even be more substantial if the ISP has to offer access to the BE class free of charge (like under the status quo).

The economic literature on QoS tiering on the CSP side is very young, but many papers have been published in recent years. Theoretical models rely on assumptions about the network

[9]Skype offers a VoIP service that allows free calls between all connected users.

industry, that may influence many of the results. First there are differences about the nature of Internet service quality. Some authors view data transportation quality as a complement to quality of the delivered content, whereas others see it as substitutes. Jamison and Hauge (2008) focus on the question whether transmission quality can substitute for content quality. They assume that the current network capacity will increase with the introduction of QoS, such that the transmission quality of the BE class is not affected. They find that QoS has a beneficial effect on content and service variety. This result stems from the fact that CSPs with a lower initial value can now compensate for that through buying priority access.

Other critical assumptions are made with respect to data delivery quality and demand for bandwidth. Economides and Hermalin (2012) analyze the negative impact of the so-called 're-congestion effect' on overall congestion under QoS tiering. The re-congestion effect describes the assumption that those CSPs that are prioritized under a QoS tiering regime will receive even more consumer requests and thus generate more traffic than under NN, which in turn re-congests the network. This assumption is in contrast to Hermalin and Katz (2007) who argue that the overall demand for capacity is independent of priority. In Economides and Hermalin (2012) the assumption about increase in demand (re-congestion) on the one hand and the assumption about fixed capacity on the other hand leads obviously to a welfare loss by construction. In addition to that much of the analysis is based on the implicit assumption that content variety is exogenous and the same under NN and QoS tiering, resulting in NN being superior in the short-run. This assumption neglects possible positive effects on content and service variety through the provision of higher quality that would be not delivered under NN. In case the re-congestion effect is not too strong QoS tiering is the more efficient regime, because it provides higher investment incentives (i. e. , investments are not overcompensated by increase in demand).

The paper of Hermalin and Katz (2007), which rests on the same principal modeling assumptions, analyzes CSP tiering by a monopolistic ISP and also in a duopoly when the ISP is free to offer a menu of qualities (NNN) and when it is restricted to offer one quality (NN). The ISP offers a menu of qualities, because he can not observe the valuation for quality of each CSP. One obvious result of NN (single quality constraint) is that all CSP with a lower valuation for quality are driven out of the market, while CSPs with a high valuation for quality are suffering from the underprovision of quality. However, there are some CSPs with a medium valuation for quality, that are now receiving a better quality than under NN.

Cheng, Bandyopadhyay, and Guo (2011) and Choi and Kim (2010) were the first to employ a queuing model to formalize the relationship between priority and BE traffic. Both papers investigate the competition of CSPs in a duopoly and the ISP's investment incentives under CSP tiering. The models account for the effect of packet queuing in data transportation on the average waiting time in a network with respect to the overall network capacity and usage of the network. This approach allows to formalize the negative effect of prioritization on the remaining BE class. The authors investigate the effect of CSP tiering on the competition of CSPs that offer similar services. In both models, exactly two competing CSPs are located at the end of a standard Hotelling line, and it is assumed that customers dislike congestion and visit one of the two CSPs exclusively. CSPs can improve their competitive position by purchasing priority access from a monopolistic ISP. This alleviates the effects of the network congestion for the customers of the respective service to some extent. However, Choi and Kim assume that the ISP sells priority access to only one of the two CSPs exclusively. Choi and Kim make this assumption in order to exclude a possible prisoners' dilemma situation that is observed in Cheng et al. (2011), who allow the ISP to sell priority transmission to both content providers. They find that when the difference in profit margins between the two content providers is rather small, both will individually buy priority access. In this situation, neither CSP gains an advantage and the price paid for priority access is forfeited. Choi and Kim assume discriminatory access to the priority class (only one CSP can buy priority) and even more restrictive, they allow the ISP to auction off the priority lane (discriminatory surcharge). In terms of the possible pricing schemes presented in this paper, this can be considered the worst case scenario of pricing for CSP tiering. Cheng et al. (2011) as well as Choi and Kim

(2010) show for a large range of parameters that the ISP's incentive to invest in infrastructure is higher under NN, whereas CSP tiering is generally welfare-enhancing in the short run.

The paper by Krämer and Wiewiorra (2012), which corresponds by and large to Chapter 3, models the main arguments of the net neutrality debate in a two-sided market framework with network congestion sensitive CSPs and IUs on each side, respectively. The platform is controlled by a monopolistic Internet service provider, who offers CSPs prioritized access to his customers for a non discriminatory surcharge. The CSPs are not in direct competition to each other, but the model allows the entry of new CSPs and can therefore account for the impact of QoS on content variety. CSP tiering functions as a means to allocate congestion away from the congestion sensitive to the congestion insensitive CSPs. Krämer and Wiewiorra (2012) find that CSP tiering may be the more efficient regime in the short run. In the long run, it provides higher incentives for broadband investments, because the entry by new, congestion sensitive CSPs creates additional demand for the priority service and consequently additional revenues for the ISP. However, long run welfare results depend on the distribution of congestion sensitivity on the CSP side of the market. If the mass of congestion sensitive CSPs is very large, an effect similar to the re-congestion effect of Economides and Hermalin (2012) can be observed. This shows how dependent welfare results are in relation to ad-hoc assumptions about the development of traffic consumption.

Reggiani and Valletti (2012) extend this approach by adding a single big CSP (e. g. , Google) to the model, which offers different services simultaneously, while a continuum of small CSPs ('the fringe') offers only one service per company. Reggiani and Valletti find that NN is likely to hinder investment at the core, but in contrast to that innovation at the fringe is stimulated. Still they can show that the service development of the big CSP is reduced. Like in Krämer and Wiewiorra (2012) CSP tiering leads to a better allocation of network resources and is consequently welfare enhancing. Nevertheless CSP tiering may eventually be beneficial to the big players on the CSPs side of the Internet ecosystem.

The question remains what policymakers and regulators could do about the potential dangers under a CSP tiering scenario, besides enforcing a NN regime and possibly risking reduced infrastructure investments in the future. They could implement a Minimum Quality Standard (MQS) to deal with this threat instead. A MQS could be introduced solely for the BE class or for each service class individually. Very often proponents argue that 'making someone faster' is of no concern to them, but only if at the same time 'no one is made worse off'. Insights from queuing theory show that this is only possible if the network capacity is increased. Otherwise the trade-off explained before still holds. The reason for this results is in direct relation to the inherent property of the $M/M/1$ queuing model, as described before. The rationale behind this argument will be further analyzed in a theoretical model in Section 3.6. Brennan (2011) argues in favor of a MQS and shows that this regulatory tool may have far less negative impact than NN. It can mitigate not only QoS related concerns, but also traffic management related concerns that relate to the artificial degradation of rival content or costly traffic flows. He also notes, however, that an MQS could also be misused by the industry. Incumbents could vote for a high quality standard in order to foreclose the market for (low quality) entrants.

2.3.2. Internet User Tiering Scenario

NN Definition 1.1 does not narrow down the implication of neutrality solely to the CSP side of the market, but applies to the transmission of data packets in general. According to this definition it would be forbidden to slow down or speed up data, independent of the side of the market requesting preferential treatment. Following this strict definition it would be a violation of NN if some users would get preferential treatment in data transportation and others would not, even if CSPs are not charged at all.

In contrast to the US, in Europe this scenario is not in the focus of the public debate. One potential reason might be the highly competitive market environment in Europe, and the fact that NN is mostly associated with discrimination of CSPs, which would only have indirect effects on consumers. Many business models in the Internet economy are based on advertisements. Therefore consumers do not fear immediate price increases through priority transmission fees that are firstly payed by the CSPs. From a business perspective differentiated service on the IU side of the market could also be an additional revenue stream that compensates the ISP for increasing infrastructure costs and mitigates the effect of costly overprovisioning in the network. Some ISPs already offer Internet access tariffs to their customers that incorporate premium options for preferential treatment. The British ISP ‘Plusnet’, for instance, offers three service classes to its customers. The service classes discriminate among different types of services as can be seen in Figure 2.2.

Figure 2.2.: Plusnet service classes (from Plusnet (2011))

To explain and justify this procedure the provider clarifies:

> “As customer numbers and usage grows, we have to increase the capacity of our network to make sure there’s enough bandwidth for everyone. When the network is busy, it’s easy for things to get out of control. Peer-to-peer, video streaming and large file downloads can flood the network quickly and use up the bandwidth. If this happens, it reduces the speed of web browsing, email, gaming and other types of web traffic. Traffic management prevents this from happening. With

> Traffic Management we can do lots of clever things to make sure everyone gets a good, fair online experience" (Plusnet, 2011).

Before the NN debate emerged, MacKie-Mason and Varian (1995) studied congestion pricing in a theoretical model as a tool to encourage users to efficiently use available network resources. They compare the equilibrium with and without usage fees and conclude that one key element of network costs is the social cost of congestion. They argue that pricing is the best way to ensure efficient network operation because it has the advantage that it "[...]plays a dual role: it provides a measure of the social cost of increased usage for a given capacity, but it also determines the value of a change in capacity" (MacKie-Mason and Varian, 1995, p.1142).

In an article that was also not inspired by the NN debate, Reitman (1991) looks into endogenous quality differentiation in congested markets. He distinguishes between impatient and patient consumers and finds that in an competitive equilibrium firms will always choose prices and capacities to differentiate themselves from each other. With many firms in the market, all of them are choosing the same level of capacity and differentiate each other solely by prices. This result is similar to the idea of Paris Metro Pricing (Odlyzko, 1999; Chau et al., 2010). Odlyzko states that until the mid 1980s the Paris Metro service was operated with first and second class cars. The cars in both classes were absolutely identical, but first class tickets cost twice as much as second class tickets. Therefore, congestion in the first class cars is lower. With respect to eyeball ISPs, Paris Metro Pricing would mean that the available bandwidth is split up between priority services and BE services. The consequence of splitting up capacity and differentiating prices is that more revenue is generated compared to a single network with the entire capacity.

In a similar model, Schwartz et al. (2008) investigate the investment and capacity division decision of competing ISPs. They also find that two service classes are socially beneficial. Moreover, the proportion of consumers that is worse off than with a single service class is decreasing in the number of ISPs in the market. If there were only a few ISPs with market power, too much capacity would be provided for the premium service and thus a larger share of consumers would be forced into this service class, resulting in a welfare loss. Therefore, the authors conclude that capacity regulation (i. e., the assignment of a predefined share of capacity to the standard service) would be welfare enhancing. This regulation would be similar to a MQS regulation (cf. Section3.6). The authors assume that users could otherwise boycott the transition to a new QoS regime. Ex-ante MQS regulation can consequently build reputation for a QoS regime and increases the probability that the transition to a new tiered system successfully takes place. Under low competition ISPs anticipate NN regulation and prefer to stay with the status quo (i. e., the ISPs assign all capacity to one service class), because regulation places uncertainty on QoS related network infrastructure investments.

Also Walrand (2008) shows under a similar setting that total welfare is higher if capacity is split up and two service classes are offered. However, he finds that total revenue would be maximized if one ISP offers the high quality (expensive) network and the other ISP the low quality (cheap) network. The higher revenue is made by the low quality ISP and therefore the ISPs face a dilemma which could lead to a price war. In this case both ISPs try to position themselves as the cheap network provider. At some point the price is so low, that it becomes profitable again to deviate and position as the expensive ISP and consequently no stable equilibrium in prices might exist.

Bandyopadhyay and Cheng (2006) study the effect of IU tiering in a theoretical model that incorporates queuing theory. In their model, users are able to decide for preferential treatment of their data on the fly during an Internet session. They find that this pricing scheme would increase the revenues of ISPs without significant costs to introduce the service. Nevertheless, the authors identify a potential problem: The monopolistic ISP might want to serve only customers with a high valuation for this service, excluding other customer groups (separating equilibrium). This could make regulatory intervention necessary provided that Internet access is politically classified as an universal access technology.

Since user side quality of service is currently not rolled out by a significant number of ISPs the question remains, why ISPs lobby so intensely for the introduction of differentiated services on the CSP side, but still hesitate to offer differentiated service to the IUs. A business strategy that seems to be less risky in terms of regulation and general juridical hurdles.

Interestingly enough, regulators and politicians are less concerned with this kind of preferential treatment. This scenario would even be in line with Definition 1.3 of Sir Tim Berners-Lee. One potential reason for this puzzle might be the perception of price discrimination (i. e., differentiated service). Fairness seems to be very important when dealing with pricing schemes for end-users. In contrast to CSPs the IUs perception of pricing schemes in general is more prone to psychological and social influence factors (Bolton et al., 2003). This effect becomes even more prominent if the good is scarce and has to be allocated between all users. Chapter 4 evaluates the aspect of fairness in Internet access pricing w. r. t. to QoS and different pricing schemes. To this end a web-based survey was conducted via an online panel among 1035 representative Internet access customers. The chapter identifies several key elements determining the consumer perception of pricing schemes and applies them to the Internet domain. To understand how users would possibly react to the introduction of QoS tariffs in general one has to find out, which of the identified factors (e. g., fairness) influence the perception of differentiated services. It is also unclear if the (level of) acceptance depends on the way how data is prioritized, i. e., capacity distributed among the customers of an ISP.

If customers feel treated unfair or perceive changes in service delivery as unfair their reactions can reach from simple complaining to a complete boycott of the product or the delivering company in general. The NN debate shows in an impressive way to what extent the complaint of Internet activists can grow.

2.4. Net Neutrality and Regulation

This section presents a policy decision guideline in form of a clustered decision process and gives a short overview over the actual regulatory actions in the US and Europe. The guideline is based on results from the NN literature and will be compared to the reasoning of real regulatory agencies.

2.4.1. Decision Support for Regulators

Any possible deviations from NN are driven by two fundamental assumptions. First, the believe that Internet traffic will increase at a rate which cannot be handled by overprovisioning and traffic management techniques and which will therefore result in a (future) congestion problem. Second, the ISPs claim they cannot afford the resulting costs for the necessary network infrastructure investments. Thus, they claim that it is inevitable to switch to a NNN scenario in the near future. If regulators agree that both assumptions are true, a switch to a NNN senario should not be prohibited ex-ante. The previous survey of the academic literature has shown that under certain NNN scenarios also gains exist, that should be considered and weighted against the inherent dangers of those NNN scenarios before deciding for any form of intervention.

Figure 2.3 shows the potential threats that are associated with the different NNN scenarios and clusters them into those that are specific to a tiered QoS system (CSP tiering and IU tiering scenario), two-sided pricing (termination fee scenario) or a managed network (status quo).

For all QoS tiering scenarios the threat of the dirt road fallacy exists. Consequently, a MQS policy could be applied to prevent artificial quality degradation or reduced capacity expansion. Only if this measure is believed to fail, direct NN intervention is appropriate.

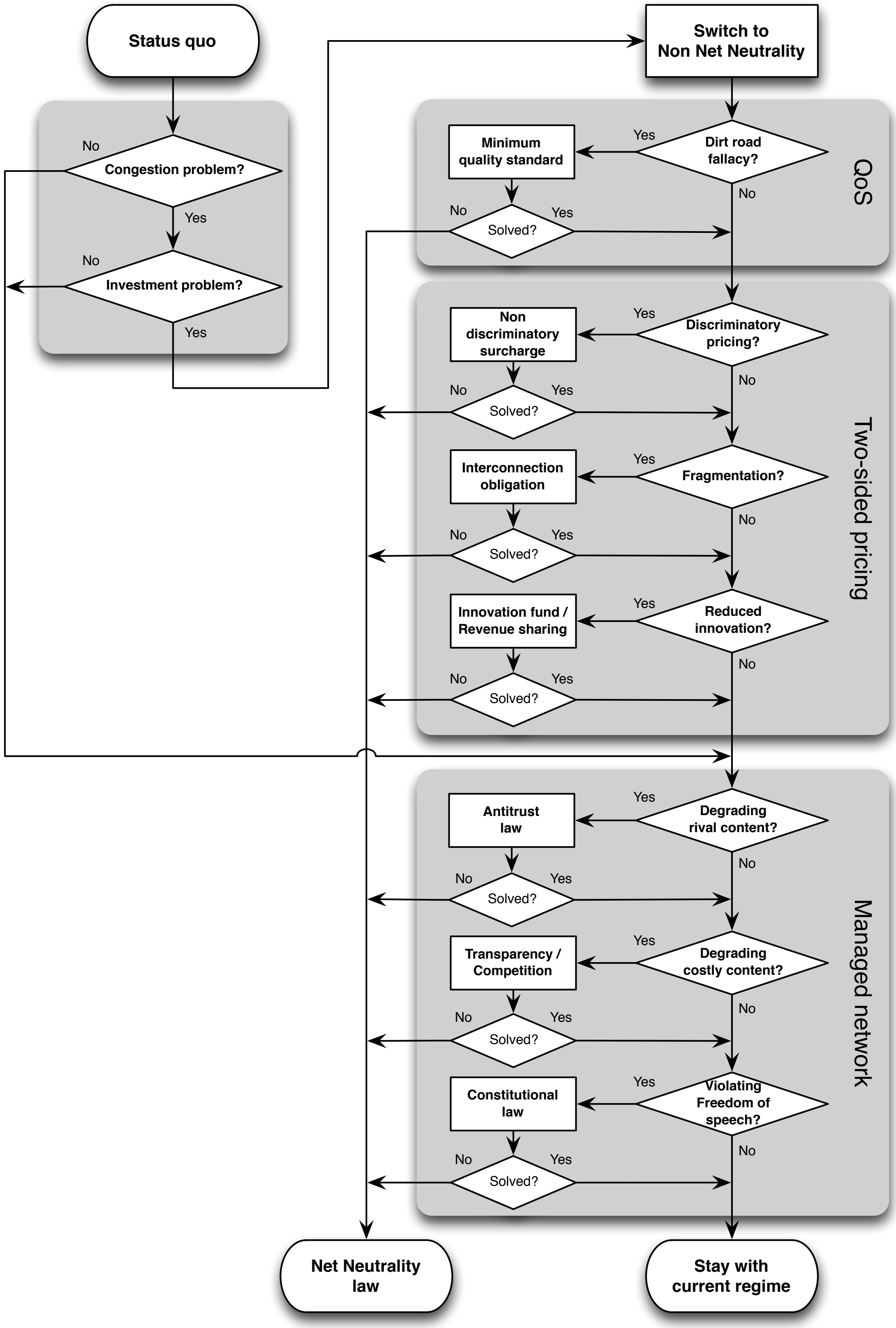

Figure 2.3.: NN policy decision process

The same logic applies to the concerns related to two-sided pricing. First, if ISPs engage in discriminatory pricing, a non discriminatory surcharge could be demanded, based on whatever criteria appear suitable (e. g. , Definition 1.4). Another concern under two-sided pricing is reduced innovation at the edge, which might result from positive termination fees that cannot be afforded by smaller CSPs. Regulators could rely on innovation funds (venture capital) or revenue sharing agreements, as described in Section 2.2.1. Lee and Wu (2009) discuss the threat of fragmentation of the Internet, which could possibly arise if ISPs and CSPs cannot agree on a suitable termination fee, or if ISPs make exclusive deals with certain CSPs. A possible remedy would be to enforce mandatory interconnection similar to the telecommunications industry. Only if one of these measures cannot solve the possible NNN dangers, stricter NN seems justified.

The final set of concerns and remedies is related to practices that can already be employed by eyeball ISPs in the current status quo, i. e. , in a managed network. Distortion of downstream competition (degradation of rival content, e. g. , VoIP) can readily be addressed ex post by antitrust law. Likewise, violations of freedom of speech are subject to constitutional law. Furthermore, if an ISP engages in degradation of certain users, content or protocols due to cost considerations, a mix of transparency obligations and competition is currently considered to be the right regulatory response. Again, only if these measures cannot address the concerns adequately, stricter NN regulation should be taken into consideration.

The proposed regulatory decision support shows that (strict) NN regulation should only be applied if the proposed remedies are believed or proven to fail. Furthermore, if a measure is believed or proven to fail, one should try to adjust or refine the measure first, to address the reason of failure. Recall that regulators also have to consider the possible efficiency gains of specific NNN scenarios and the implementation costs of each remedy they intend to apply. However, it seems that already many tools under the status quo exist that can deal with the most acute threats.

2.4.2. Regulation in the United States and Europe

In 2005 the FCC adopted four principles to ensure the 'open and interconnected character of the Internet', a circumscription to avoid the biased term NN. In the Report & Order from 2010 the FCC proposed an updated version of this goals with a stronger focus on transparency.

> "A person engaged in the provision of fixed broadband Internet access service, insofar as such person is so engaged, shall[...]"
>
> 1. Transparency
>   "[...]publicly disclose accurate information regarding the network management practices, performance, and commercial terms of its broadband Internet access services sufficient for consumers to make informed choices regarding use of such services and for content, application, service, and device providers to develop, market and maintain Internet offerings."(FCC, 2010, §54)
> 2. No Blocking
>   "[...]not block lawful content, applications, services, or non-harmful devices, subject to reasonable network management."(FCC, 2010, §63)
> 3. No unreasonable degradation
>   '[...]not unreasonably discriminate in transmitting lawful network traffic over a consumer's broadband Internet access service."(FCC, 2010, §68)
> 4. Reasonable network management
>   "A network management practice is reasonable if it is appropriate and tailored to achieving a legitimate network management purpose, taking into account the particular network architecture and technology of the broadband Internet access service."(FCC, 2010, §82) [10]

[10] *"Legitimate network management purposes include: ensuring network security and integrity, including by addressing traffic that is harmful to the network; addressing traffic that is unwanted by end users (including by premise operators), such as by providing services or capabilities consistent with an end user's choices*

This shows that the FCC is more concerned with regulating the status quo. As exemplified in Section 2.1.1, transparency regulation makes only sense in a situation with sufficient competition. However, compared to the 2005 principles, which explicitly mentions the 'right of consumers for competition among network providers', the current Report & Order relies predominantly on transparency. Nevertheless, it is unlikely that any deviation to a QoS regime will be supported by the FCC since pay for priority arrangements would "raise significant cause for concern" (FCC, 2010, §76). It is interesting to note that the ruling excludes all wireless network providers. The FCC argues that there is sufficient competition between wireless operators and that compared to the fixed line business it is still an emerging market. For this reason and in line with its own definition of NN (Definition 1.4), Verizon could introduce QoS in its network.

In the European Union (EU) no specific NN law exists either. The EU commission adopted a 'wait and see' approach and is less concerned about the introduction of a possible CSP tiering regime (specifically, EU Directive 2009/136/EC from November 25, 2009). However, the directive relies on similar transparency and non-blocking rules as the FCC's order (FCC, 2010). In contrast to the US the EU allows the national regulatory agencies to set a MQS to prevent undesired actions of ISPs under the status quo or a potential dirt road fallacy (cf. Article 22, 3).

Besides these general regulatory approaches to NN in the US and the EU, there exist two notable national exceptions.

First, Chile enacted a NN law in 2010, which was finally implemented in May 2011 (Art.24Ha/Ley 20.453). At first, the law was considered to be the only implementation of strict NN in the world (Gaitonde, 2010). However, the final amended regulation states that ISPs cannot "arbitrarily block, interfere, discriminate, hinder or restrict" the use of the Internet. This passage reveals that the law prevents just arbitrary discriminatory actions and not a tiered system per se. In this respect the final law is more a compromise between NN propo-

regarding parental controls or security capabilities; and reducing or mitigating the effects of congestion on the network."(FCC, 2010, p.17952).

nents and opponents than a codification of strict NN (Nixon, 2011).
Second, the Netherlands enacted a NN law in 2011 that forbids network operators to degrade certain applications or charge extra fees to its customers. The law was suggested after an announcement of mobile network operators to charge extra for certain VoIP and messaging applications. Being the first country in the EU enacting a NN law, the Netherlands followed the example of Chile, but in contrast to the US the law targets the emerging mobile markets (O'Brien, 2011). In May 2012 the law is expected to be finally implemented by the Dutch senate.

In Germany, the government commissioned a committee of inquiry (partially comprising of politicians and experts) on different issues of the digital economy, among which was also the issue of NN. In its final report, which is also intended to serve as a guidance to the national regulatory authority among which NN was also discussed, the committee did not come to a consensus w. r. t. CSP tiering. However, similar to the FCC's order, it acknowledges that reasonable network management is welfare enhancing.

Chapter 3.

Content and Service Provider Tiering

"Now what [content providers] would like to do is use my pipes free, but I ain't going to let them do that because we have spent this capital and we have to have a return on it."

Ed Whitacre (Chief Executive Officer of ATT) (O'Connell, 2005)

THIS chapter is intended to extend the existing theoretical literature about NN that uses queuing theory to analyze a possible CSP tiering scenario. The chapter corresponds in large parts to the paper by Krämer and Wiewiorra (2012).

3.1. Motivation

In the ongoing NN debate CSP tiering has been identified as the most desired regime by ISPs. However, current models relying on queuing theory predominantly explore the aspect of competition between two competing CSPs, that have been discussed in Section 2.3.1. Due to the

Hotelling (1929) framework, these models have several drawbacks. First, they assume that IUs use only one of the two services exclusively (single-homing). Consumers using Google's search engine would consequently never use Microsoft's Bing search, too. Especially in the context of the Internet this assumption seems to be very restrictive and unrealistic. NN proponents are particularly concerned with the variety of content and services in the Internet ecosystem because IUs have a valuation for variety. This essential aspect of the debate can not be addressed by this class of models. This drawback is closely related to another problematic restriction. Both models that make use of results found by queuing theory (cf. Table 3.1) assume only a fixed number of CSPs (i. e. , two). Although this captures the notion of competition sufficiently, it does not allow to study the impact of QoS on content variety. Consequently, with this class of models, one is not able to analyze whether CSPs are driven out of the market in case they cannot afford additional access fees. Therefore it is also not possible to make predictions about the relation between network investments and content variety. As exemplified in Chapter 1, ISPs claim that only a shift to a new network regime ensures that they can cope with the increasing demand for bandwidth. Second, existing models neglect the effect of congestion sensitivity. Section 2.3 gave an introduction into the relation of neutrality in data transportation and the demand for quality of different types of services and content. Recall the simple fact, that e. g. , real time applications suffer more from congestion than e. g. , a simple mail service.

As described in Section 2.2, the Internet can be modeled as a two-sided market, with CSPs and Internet customers on either side, each of which value an increasing number of actors on the other side and dislike network congestion. It is assumed that the ISP has a terminating monopoly over its customers (e.g., due to the customers' lack of alternative ISPs or high switching costs). Although this may sound restrictive at first, this is a reasonable assumption for many regions in the US and Europe. For instance in the US many regions are covered only by the local cable company. Therefore, the only way for the CSPs to reach those customers is through the ISP's network. Although the CSPs' customer base is probably comprised of customers of many different ISPs, each of which might have a terminating monopoly, it is still insightful to investigate the relationship between CSPs and a single ISP, particularly if that ISP is thought to be large. For example, it would certainly have a substantial impact on CSPs' business model if they would not have access to customers' on AT&T's network in the United States of America (USA), or the customers of German Telecom.

Note that the following model deals only with one of the NNN scenarios presented in Section 2, namely CSP tiering. In contrast to e. g., Choi and Kim (2010) the model is designed to analyze a more realistic pricing structure. As explained in Section 2.4, one of the inherent dangers of two-sided pricing is a discriminatory surcharge, i. e., CSPs are charged different prices even if they are in the same class of services (e. g., VoIP). It is very likely that regulators would never accept a shift to a QoS tiering regime, especially not a CSP tiering scenario, if non-discrimination is not warranted. To study the effect of CSP tiering without the inherent drawback of discrimination, the following model considers an ISP that charges only the same price per transmitted data packet to all CSPs buying priority access.

Another concern mentioned in Section 2.4 is reduced innovation and fragmentation of the Internet. To exclude this threat ex-ante the model only considers those charges to the CSPs that are over and beyond those for access to the Internet. Consequently, the BE lane under CSP tiering remains to be offered for zero additional costs. Analog to that, NN is considered as a zero price rule which implies that the ISP cannot charge CSPs additionally for terminating traffic in its network.

However, the model does not assume any restrictions on the price for priority access, nor the quality that is offered under QoS tiering. Therefore it is not clear ex-ante if a switch to this non-discriminatory CSP tiering regime is welfare enhancing. However, it would be highly unlikely that CSP tiering without these restrictions would ever be allowed by regulators, following the argumentation of e. g., the FCC (cf. Section 2.4). Consequently, these restrictions are restrictive in appearance only, but are in fact realistic assumptions to model a benchmark for a possible CSP tiering scenario. Conversely it would be wrong to generalize the following findings to any possible forms of CSP tiering: However, the model presented in this Chapter allows to analyze the following research questions.

Research Question 1. *What is the effect of a non-discriminatory CSP tiering scenario in a local access monopoly in the short-run?*

> **Research Question 1.1.** *What is the effect of a non-discriminatory CSP tiering scenario in a local access monopoly on content variety?*
>
> **Research Question 1.2.** *What is the effect of a non-discriminatory CSP tiering scenario in a local access monopoly on CSP, IU and ISP welfare?*

Research Question 2. *What is the effect of a non-discriminatory CSP tiering scenario in a local access monopoly in the long-run?*

> **Research Question 2.1.** *What is the effect of a non-discriminatory CSP tiering scenario in a local access monopoly on capacity investments?*
>
> **Research Question 2.2.** *What is the effect of a non-discriminatory CSP tiering scenario in a local access monopoly on content variety?*
>
> **Research Question 2.3.** *What is the effect of a non-discriminatory CSP tiering scenario in a local access monopoly on CSP, IU and ISP welfare?*

Research Question 3. *What is the effect of MQS regulation under a non-discriminatory CSP tiering regime in a local access monopoly?*

Table 3.1 shows a summary of the important features of the model in comparison to the relevant previous literature in the field.

Table 3.1.: Comparison of related work

	Non discriminatory surcharge	**Infrastructure investments**	**Entry of CSPs**	**Congestion Sensitivity**
Choi and Kim (2010)	□	■	□	□
Cheng et al. (2011)	■	■	□	□
Krämer et al. (2012)	■	■	■	■

3.2. The Model

3.2.1. Content and Service Providers

The model considers a continuum of CSPs. This allows for an explicit analysis of the effect of a CSP tiering scenario on content and service variety in the Internet ecosystem. Whatever service the CSPs offer, they generate revenues only indirectly through online advertisements. This rules out the possibility of CSPs charging consumers directly for access to their content. Although this may sound restrictive at first, it has been exemplified in Section 1.4.1 that this is the more relevant case, as empirical evidence suggest that customers are generally fairly reluctant to pay extra for specific content or services (Dou, 2004; Sydell, 2007).

In the model, a CSP's advertisement revenue will depend on the average received traffic, the per-click advertisement revenue and its unique click-through rate. The click-through-rate of each CSP is determined by its innate sensitivity towards network congestion. This allows to explicitly account for differences in congestion sensitivity between content and services in the Internet that have been explained in Section 1.4.3.

Before these measures are formally introduced below, one fundamental assumption has to be made:

Assumption 1. *Each CSP receives the same average traffic from each customer, denoted by* λ*. This is independent of a CSP's business model and consequently its innate sensitivity to network congestion.*

It will often be convenient to think of λ as the number of 'clicks' that a customer generates on each CSP's website. This assumption provides a neutral and unbiased reference case w. r. t. the traffic that is generated by the specific CSPs and w. r. t. the value of the individual content of the CSPs. This assumption does therefore not imply any ex-ante re-congestion effect. It has already be discussed that e. g. , Economides and Hermalin (2012) assume an increasing traffic for CSPs that are connected with higher quality. However, this obviously would introduce an ex-ante disadvantage of CSP tiering that has been shown to drive the results of such models (cf. Section 2.3.1).

In general the relationship between congestion sensitivity of a CSP and the amount of traffic that this CSP generates is far from obvious. For example, VoIP services are highly congestion sensitive (in terms of jitter, delay, packet loss), but generate comparably little traffic. Likewise, file hosting services are highly traffic intensive but not congestion sensitive. Assumption 1 is aimed at avoiding this kind of relationship between the traffic that a CSP generates and its congestion sensitivity, since it would also bias the analysis. Consequently, this setup allows to assess CSP tiering based on its core ability, i.e., increasing transmission quality (not bandwidth). Moreover, by applying Assumption 1 any judgment about the value of specific content or services to consumers is avoided. In the model, CSPs offer heterogeneous services which are all equally valuable to customers. Therefore it is reasonable to assume that customers distribute their clicks evenly among the available CSPs. As explained before it is not intended to study the effect of CSP tiering on the direct competition between otherwise similar CSPs. Therefore, Assumption 1 implies that the model abstracts from any business stealing effects. More specifically, in what follows, it is assumed that λ is constant and thus, as the number of active CSPs increases, consumers increase their total number of clicks accordingly. Competitive pressure is introduced through diminishing ad revenues, as described below.

Nevertheless it has to be controlled for the alternate assumption, that the consumers' total number of clicks is fixed and thus λ diminishes as more CSPs enter the market. This would introduce a general notion of competitive pressure among the CSPs, i.e., as more CSPs enter the market, the revenue of each CSP is reduced, keeping everything else constant. This alternative setup is analyzed in Section 3.5.1. Note that the results of either assumption are qualitatively the same.

Congestion sensitivity on the CSP's end has the effect that only a fraction of the arriving clicks can be turned into advertisement revenue. This measure is known as the click-through rate. It is assumed that each CSP's click-through rate diminishes as network congestion increases. Moreover, CSPs are heterogeneous w. r. t. to their sensitivity towards network congestion. Each CSP's business model has an innate sensitivity as to what extent network congestion affects the click-through rate. For example, a web-based e-mail provider is likely to be relatively insensitive to network congestion. Consumers that arrive on the website are satisfied with the service even under high network congestion, and more likely to click on advertisements, because their QoE with the service is still satisfactory. In contrast, consumers of a highly congestion sensitive web service (e.g., online gaming) may still sign up for the CSP's gaming service, but are less satisfied in the presence of network congestion. The QoE with the service is worse and therefore IUs are less likely to click on advertisements. This individual congestion sensitivity is denoted by θ and the corresponding click-through rate of a CSP is assumed to be $(1-\theta w)$, where w denotes the average waiting time in the network, the indirect measure of the level of network congestion. The click-through rate itself follows a so-called 'Poisson thinning process'. The probability that an arriving IU becomes advertising relevant to the CSP depends on the average waiting time in a service class (w) and the sensitivity of the service (θ) itself. Therefore a CSP with a high innate sensitivity has a lower probability of making money from the same amount of arriving traffic than a CSP with a low innate sensitivity at any given congestion level.

In the following analysis it will be distinguished between the NN regime and the QoS tiering regime (CSP tiering scenario) by subscript N and Q, respectively[1]. Assume that there exists a continuum of CSPs with unit mass and distribution function $F(\theta):[0,1]\rightarrow[0,1]$. Let r be the average revenue-per-click on advertisements depending on the mass of active CSPs in the market. Therefore, each CSP's profit under NN is

$$\Gamma_N(\theta)=\begin{cases}(1-\theta\, w_N)\,\lambda\bar{\eta}r & \text{if active}\\ 0 & \text{otherwise,}\end{cases} \tag{3.1}$$

[1]In order to simplify the notation, the subscripts will be omitted wherever the referenced network regime is unambiguous.

where $\bar{\eta}$ denotes the share of Internet customers in equilibrium. Under NN all CSPs perceive the same level of congestion, w_N.

In a CSP tiering scenario, CSPs can now opt for the priority transmission class with $w_{Q1} < w_N$ at a price of p per click. The CSPs that remain in the BE class, on the other hand, experience a higher congestion level $w_{Q2} > w_N$.

$$\Gamma_Q(\theta) = \begin{cases} (1-\theta\ w_{Q2})\,\lambda\bar{\eta}r & \text{if active in best-effort class} \\ (1-\theta\ w_{Q1})\,\lambda\bar{\eta}r - \lambda\bar{\eta}p & \text{if active in priority class} \\ 0 & \text{otherwise.} \end{cases} \tag{3.2}$$

The CSP that is indifferent between choosing the priority and the BE transmission class under a QoS tiering regime, is denoted by $\tilde{\theta}$. Furthermore, in both regimes, the CSP that is indifferent between becoming active and staying out of the market is characterized by a congestion sensitivity of $\bar{\theta}$. Thus $F(\bar{\theta})$ reflects the mass of all active CSPs (content variety) and the share of CSPs choosing the priority class under a CSP tiering scenario is given by $\beta \equiv 1 - F(\tilde{\theta})/F(\bar{\theta})$.

The advertisement market is assumed to be competitive. This means that the average advertisement revenue per click depends on the mass of active CSPs, i. e., $r(F(\bar{\theta}))$, and that $\partial r(\cdot)/\partial F(\bar{\theta}) \leq 0$. This introduces some kind of indirect competition between the CSPs. The intuition behind this assumption would be a fixed budget in the advertising industry. Following Assumption 1 the overall traffic in the Internet economy increases with the entry of new CSPs. As a result of this traffic expansion the number of ad-relevant clicks also increases, which in turn leads to a lower advertisement revenue per click for all CSPs since the overall budget is fix.

3.2.2. Customers

IU value basic connectedness to the Internet as well as the presence of many CSPs on the other side of the market. On the one hand, basic connectedness adds a fixed utility of $b > 0$ whereas each additional CSP adds a marginal utility of $v > 0$ to a customer's utility.[2] On the other hand, congestion reduces a customer's utility of using the CSPs' services.

It is assumed that consumers' utility is diminished by the average congestion level in the network $w_Q = \beta w_{Q1} + (1-\beta)\, w_{Q2}$, or w_N, respectively. This implies that $w_Q = w_N$ in the short run (when the capacity of the network is fix and consequently equal in both regimes, i. e., $\mu_Q = \mu_N$) whenever the same amount of CSPs enters the market (i. e., $\bar{\theta}_N = \bar{\theta}_Q$).
Alternatively, one could have assumed that customers are congestion sensitivity aware w. r. t. each individual CSP and that they instead evaluate the level of congestion as $\hat{w}_N = \int_{\theta=0}^{\bar{\theta}} w_N \theta f(\theta) d\theta$ and $\hat{w}_Q = \int_{\theta=0}^{\tilde{\theta}} w_{Q2} \theta f(\theta) d\theta + \int_{\theta=\tilde{\theta}}^{\bar{\theta}} w_{Q1} \theta f(\theta) d\theta$, respectively. This alternative model variant will be analyzed in Section 3.5.2.

For the reminder of this chapter it will be assumed that consumer's utility depends on the average level of congestion. This assumption does not qualitatively change the analysis, but makes the analytical solutions clearer. If consumers additionally evaluate congestion sensitivity based on the sensitivity parameter of the connected CSPs, CSP tiering allocates congestion more efficiently such that $\hat{w}_Q \leq \hat{w}_N$. Therefore, abandoning this assumption emphasizes the potential advantages of the CSP tiering scenario. The analysis is consequently analytically better traceable, but also more conservative w. r. t. QoS, and tipped in favor of the NN regime.

[2]This assumption is made to avoid any presumptions about the relation between congestion sensitivity and the value of a service. Again, this is in contrast to e. g. , Economides and Hermalin (2012), but can be considered as an unbiased reference case.

Formally,

$$U = \begin{cases} b + v\bar{\theta} - \iota w - a & \text{if connected} \\ 0 & \text{otherwise,} \end{cases} \tag{3.3}$$

where $\iota > 0$ denotes a consumer's marginal disutility due to congestion, and a the Internet access fee charged by the ISP.

As outlined before, the analysis focuses on the effect of the CSP tiering scenario under monopoly. Thus, for expositional clarity, it is assumed that customers are homogeneous and that therefore the ISP is able to set an access fee such that all IUs connect to the ISP in equilibrium. It is very important to note, that this assumption does *not* violate the two-sided market property (cp. Rochet and Tirole, 2006) and is not a crucial limitation of the model per se.

With heterogeneous customers this fee is likely to be higher under NN than under CSP tiering where the ISP can collect additional rents from CSPs through the priority fee. Under CSP tiering the ISP may even find it profitable to subsidize the consumer side by lowering the access fee, possibly down to zero, in order to stimulate customer subscriptions which in turn allows the ISP to make higher profits on the CSPs' side. Consequently, the customer access fee and associated dead weight loss is lower under CSP tiering. Indeed, the dead weight loss may even be zero if the access fee is zero. Under NN, the consumers' access fee (and thus the dead weight loss) can never be zero, because otherwise the ISP would not make any profit. Consequently, it is likely that more consumers will subscribe under CSP tiering, and that consumers' welfare is higher than under NN. This result has already been discussed briefly in Section 2.2.1 in connection with the paper of Economides and Tag (2012). In a two-sided market the price on each side depends on the valuation for the other side of the market. As long as CSPs indeed value an additional IU more than the IUs value an additional CSP, the ISP has an incentive to attract additional customers to charge a higher price on the CSP side of the market. If the logic holds the other way around, the ISP may want to increase content and service variety to attract additional customers.

3.2.3. Network Congestion

Network congestion in the model is measured indirectly through Internet consumers' average waiting time following a content request. The model makes use of the well-known $M/M/1$ queuing model (Kleinrock, 1976) to capture the relationship between average waiting time, network traffic and capacity. The basic logic behind this queuing model has been already briefly discussed in Section 2.3. Recall that queuing models in their general notation denote *lambda* as the 'average arrival rate' (in the context of this model interpreted as the *network traffic of a single CSP*) and *mu* as the 'average service rate'[3] (in the context of this model interpreted as the *network capacity*). Further, the $M/M/1$ queuing model assumes that

1. service requests arrive according to a Poisson process (i.e., arrivals happen continuously and independently of one another),

2. service time is exponentially distributed (i.e., request coming from a Poisson process are handled at a constant average rate) and

3. that service requests are processed by a single server.

This last assumption is equivalent to assuming that network performance is dominated by a bottleneck component. Furthermore it is assumed that the length of the queue as well as the number of users is potentially infinite. This model is standard and considered to be a good proxy for actual Internet congestion (McDysan, 1999).

Under NN the $M/M/1$ model predicts that each consumer has an expected average waiting time of

$$w_N = \frac{1}{\mu - \Lambda}. \tag{3.4}$$

[3]The speed at which the server handles the arriving data packets.

$\Lambda = \lambda\ \bar{\eta}\ F(\bar{\theta})$ denotes the average rate at which customers' aggregate content requests arrive at the ISP's network, which is interpreted as the *overall network traffic*. For the queuing system to be stable, it has to be assumed that $\mu > \Lambda$.

Under a CSP tiering regime, CSPs are offered the choice between a priority and a BE transmission class. In the $M/M/1$ model this translates into introducing an additional queue which handles the request of the CSPs in the priority class and which is processed ahead of the queue for the BE class. However, in each class the queue is cleared on a first-come first-served basis. In this vein, the classical results of the $M/M/1$ queuing model represent the average waiting time in the priority class, w_{Q1}, and the BE class, w_{Q2}:

$$w_{Q1} = \frac{1}{\mu - \beta\Lambda}, \qquad w_{Q2} = \frac{\mu}{\mu - \Lambda} w_{Q1} \tag{3.5}$$

Observe that the relation $w_{Q1} < w_N < w_{Q2}$ is always fulfilled, assuming a fixed transmission capacity $\mu = \mu_Q = \mu_N$ and $\beta < 1$. In case all CSPs are in the first priority class ($\beta = 1$), the model yields $w_{Q1} = w_N$. This is an essential feature of the $M/M/1$ model, because it shows formally that serving some CSPs with priority will (in the short run) unambiguously lead to a *degradation* of service quality for the remaining CSPs in the BE class. With respect to the NN debate and the inherent dangers of the CSP tiering scenario, recall the threat of the 'dirt road fallacy' from Section 2.3. One could argue, that this is already an indication that QoS has an detrimental effect that should be avoided. However, this argument is short-sighted. The dirt road fallacy describes the threat of *artificial* quality degradation (be it sabotage under a fixed level of capacity or intentionally reduced capacity expansion) to curb the profits from priority sales. The $M/M/1$ related degradation of the BE class is solely a result of a given capacity and the endogenous split of CSPs between the two service classes. Without sabotage the reduction in quality of the BE class is the natural result of giving priority access to the bottleneck facility. The case of artificial quality degradation will be discussed in the conclusion of this chapter.

3.2.4. Internet Service Provider

The ISP is the platform operator in the local Internet access market, over which it has a terminating monopoly. It maximizes his profit through a number of strategic variables. In the short run, it charges only an access fee, a, from connected consumers. Under NN, this is the only source of revenue for the ISP. In the long run the ISP can additionally invest into network capacity, μ. As outlined before, customers as well as CSPs dislike network congestion. The level of network congestion is captured by customers' average waiting time for content, w, which is again controlled by the ISP through its choice of network capacity.

Hence, under a NN regime, the ISP's profit is

$$\Pi_N = \bar{\eta} a - c(\mu), \tag{3.6}$$

where $c(\mu)$ denotes the costs of capacity expansion. To ensure the existence of an interior solution to the ISP's investment decision, assume a non-concave cost function, i.e. $\partial c / \partial \mu \geq 0$ and $\partial^2 c / \partial \mu^2 \geq 0$.

Under a CSP tiering regime, the ISP has an additional strategic variable, p, the price which it charges CSPs to transmit data packets with priority. The ISP will choose p in order to maximize its additional revenues from selling priority access.

More precisely, under QoS tiering the ISP's profit function is

$$\Pi_Q = \bar{\eta} a + \beta \Lambda p - c(\mu). \tag{3.7}$$

The ISP's previous investments in transmission capacity are assumed to be sunk in all regimes. Therefore, in the short run μ can be considered an exogenous variable which is irrelevant for profit maximization.

Figure 3.1 shows that this model setup is a simplified representation of the revenue streams in the Internet as presented in Section 1.4.

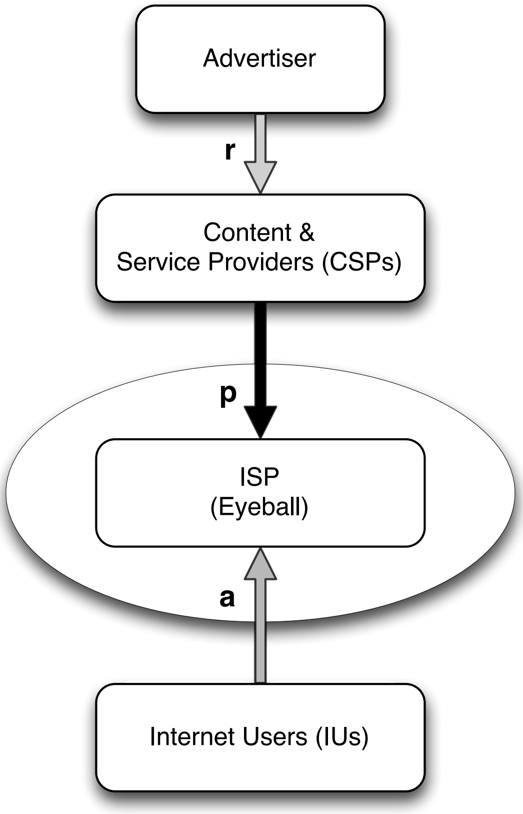

Figure 3.1.: Model CSP tiering

3.3. Short Run Effects on Content Variety and Welfare

In order to answer Research Question 1, one has to compare NN and CSP tiering when network capacity, μ, is exogenous and equal in both regimes.

3.3.1. Short Run Equilibrium and Content Variety

Net Neutrality Regime

Consider the profit maximization problem of the ISP. It is easy to see that it will set an optimal customer access charge of $a = b + v\bar{\theta} - \iota w_N$. All customers will connect to the network ($\bar{\eta} = 1$) because of the homogeneity assumption. Therefore the total consumer surplus is appropriated by the ISP.

CSPs will only enter the market if they can make a positive profit. Under NN all CSPs expect the same congestion level of w_N. Consequently, the last CSP to enter the network is located at:

$$\bar{\theta}_N = \frac{1}{w_N} = \mu - \lambda F(\bar{\theta}) \tag{3.8}$$

The analysis is restricted to the case in which (at least) the most congestion sensitive CSP, located at $\theta = 1$, remains inactive in equilibrium. This is ensured iff the average congestion level satisfies $w_N > 1$. Without this assumption the market for content and services would be already covered under NN. In this case it would be impossible to analyze a possible variety expansion effect through the introduction of QoS.

Note that an increase in network traffic per CSP, λ, has an adverse effect on network congestion ($\partial w_N / \partial \lambda > 0$) and content variety ($\partial \bar{\theta}_N / \partial \lambda < 0$). This is central to the debate on NN, because it exemplifies the network operators' concerns with respect to the expected increase in traffic.

Quality of Service Tiering Regime

Under a CSP tiering regime the CSPs have the possibility to choose between priority transmission and BE transmission of their data. Congestion sensitive CSPs are now able to alleviate the negative effect of congestion on their business by opting in the priority lane of the ISP. Obviously, each CSP has to make this decision depending on its unique level of congestion sensitivity.

Consequently, one can distinguish between three types of CSPs:

1. CSPs whose business model is relatively insensitive to network congestion. They will remain in the free-of-charge BE class.

2. CSPs whose business model is sufficiently sensitive to network congestion. They will opt for priority access at a price of p.

3. CSPs whose business model is extremely sensitive to network congestion. They will remain inactive as entry is not profitable.

Remember that the CSP indifferent between the first two cases is denoted by $\tilde{\theta}$, whereas the CSP indifferent between the last two cases is denoted by $\bar{\theta}_Q$. Obviously, it must hold that $0 \leq \tilde{\theta} \leq \bar{\theta}_Q$.[4]

Proof. Recall $\bar{\theta}_Q$ from (3.11) and notice that

$$\tilde{\theta} = \frac{p\left(\mu - \lambda(F(\bar{\theta}_Q)\right)}{r\lambda F(\bar{\theta}_Q)}\left(\mu - \lambda\left(F(\bar{\theta}_Q) - F(\tilde{\theta})\right)\right). \tag{3.9}$$

[4]Note that the model is not able to account for a business critical effect of (artificial) quality degradation on congestion insensitive CSPs. In this model the business of the most insensitive CSP is completely 'immune' to congestion ($\theta = 0$). Therefore one is not able to take into account the case, in which even the most congestion insensitive CSP would need a minimum quality to establish a sustainable business.

It follows that

$$\tilde{\theta} \leq \bar{\theta}_Q \quad \Leftrightarrow \quad p \leq r\frac{\lambda F(\bar{\theta}_Q)}{\mu}. \tag{3.10}$$

Under Assumption 2, where p is determined by (3.16) and $F(\bar{\theta}_Q)$ by (3.14) this condition becomes $\sqrt{\lambda+1} \leq \lambda+1$, which is always true. □

In a fulfilled expectations equilibrium, the last CSP to enter is located at

$$\bar{\theta}_Q = \frac{1-p/r}{w_{Q1}} = \frac{r-p}{r}\left(\mu - \lambda\left(F(\bar{\theta}_Q) - F(\tilde{\theta})\right)\right). \tag{3.11}$$

From

$$\frac{\partial \bar{\theta}_Q}{\partial p} = \underbrace{-\frac{1}{r}\left(\mu - \lambda\left(F(\bar{\theta}_Q) - F(\tilde{\theta})\right)\right)}_{\text{First Order Effect}} + \underbrace{\lambda\frac{r-p}{r}\left(\frac{\partial F(\tilde{\theta})}{\partial p}\frac{\partial \tilde{\theta}}{\partial p} - \frac{\partial F(\bar{\theta}_Q)}{\partial p}\frac{\partial \bar{\theta}}{\partial p}\right)}_{\text{Second Order Effect}} \tag{3.12}$$

one can observe that an increase in the price for priority transmission, p, has an unambiguously negative first order effect on content variety. This is the central concern of NN proponents, who argue that starting from a zero price under NN, the introduction of a positive price under CSP tiering has negative first order effects on content variety. However, this argument neglects that there is a second order effect as well: An increase in p will induce more CSPs to choose the free BE class and therefore alleviate congestion in the priority class. This in turn may encourage new, congestion sensitive CSPs to enter, which drives congestion in the priority class up again. The size and direction of the second order effect hinges on the mass of CSPs that is located locally at $\tilde{\theta}$ and $\bar{\theta}$ (i.e., $\frac{\partial F(\tilde{\theta})}{\partial p}\frac{\partial \tilde{\theta}}{\partial p} - \frac{\partial F(\bar{\theta}_Q)}{\partial p}\frac{\partial \bar{\theta}}{\partial p}$) and cannot be determined more specifically for a general distribution function.

To analyze the notion of non-uniformly distributed congestion sensitivity one has to make a simplifying assumption about the underlying density function. For the purpose of this analysis, therefore assume a particular density function of θ that exemplifies the effect of having a non-uniform distribution of θ, without loosing analytical traceability.

To this end, consider the density function $f : [0,1] \to [0,1], f(\theta) := \alpha + 2\theta(1-\alpha)$, with $\alpha \in [0,2]$. Let F be the distribution function to f and notice that for $\alpha = 1$ one obtains a uniform distribution with $F(\theta) = \theta$.
Otherwise, if $\alpha > 1$, there is a relatively larger mass of congestion *insensitive* CSPs ($F(\theta) > \theta$) and if $\alpha < 1$ there is a relatively larger mass of congestion *sensitive* CSPs ($F(\theta) < \theta$). A variation of α therefore allows one to gradually shift mass from the congestion sensitive portion ($\theta > 0.5$) to the congestion insensitive portion ($\theta < 0.5$) of the CSPs and vice versa.

It is very important to note, that under the (unbiased) assumption of uniformly distributed congestion sensitivity (i. e., $\alpha = 1$), the first order and second order effect are exactly offset. This result holds at any price level, such that the price for priority transmission has no effect on content variety under CSP tiering.

Thus, under a uniform distribution, NN and CSP tiering will yield exactly the same level of content variety, i.e., $\bar{\theta}_Q = \bar{\theta}_N = \mu/(\lambda+1) = 1/w_N$. However, if the mass of congestion sensitive CSPs is relatively large ($\alpha < 1$), an increase in price for priority will not lead to an equally large congestion alleviation for the priority class, such that the first order effect prevails. Consequently, under CSP tiering less CSPs will enter in equilibrium than under NN. Conversely, if the mass of congestion sensitive CSPs is comparably small ($\alpha > 1$), the second order effect dominates, and CSP tiering leads to more content variety than NN. The following proposition summarizes the results w. r. t. Research Question 1.1.

Proposition 1 (Content Variety). *If CSPs congestion sensitivity is uniformly distributed, CSP tiering has no effect on content variety in the short run: The number of active CSPs is the same as under NN. In both regimes the number of active CSPs is inversely proportional to the* *average level of congestion in the network. However, if the mass of congestion sensitive CSPs is comparably small (large), then CSP tiering is likely to lead to more (less) content variety.*

Proof. It is easy to verify that for $\alpha = 1$

$$F(\bar{\theta}_N) = \bar{\theta}_N = \frac{\mu}{\lambda+1} \tag{3.13}$$

$$F(\bar{\theta}_Q) = \bar{\theta}_Q = \frac{\mu}{\lambda+1} \tag{3.14}$$

$$F(\tilde{\theta}) = \tilde{\theta} = \frac{p}{\lambda(r-p)}\bar{\theta}_Q, \tag{3.15}$$

which proves the first part of the proposition. Furthermore from (3.17) it follows that

$$p = \left(1 - \sqrt{\frac{1}{\lambda+1}}\right) r = \left(1 - \sqrt{\frac{\bar{\theta}_Q}{\mu}}\right) r\,. \tag{3.16}$$

For $\alpha \neq 1$, $\bar{\theta}_Q$ will generally depend on p. First see that p cannot exceed p_{max} which solves $\Gamma_Q(\bar{\theta}_Q) = \Gamma_Q(\tilde{\theta})$. This price is given by

$$p_{max} = r\left(1 + \frac{\lambda\alpha + 1 - \sqrt{(\lambda\alpha+1)^2 + 4\mu\lambda(1-\alpha)}}{2\mu\lambda(1-\alpha)}\right)$$

. The feasible values of $F(\bar{\theta}_Q)$ and $F(\bar{\theta}_N)$ in the interval $p \in [0, p_{max}]$ are plotted in Figure 3.2. It can be readily seen that $F(\bar{\theta}_N) = F(\bar{\theta}_Q)$ for $\alpha = 1$, irrespective of the value of p, and for $\alpha \neq 1$ whenever $p = 0$ or $p = p_{max}$. In all other cases $F(\bar{\theta}_N) \neq F(\bar{\theta}_Q)$ according to the proposition.

□

Therefore, it is useful to assume a uniform distribution of θ as the reference case for the subsequent analysis, from which it is then easy to draw more general conclusions.

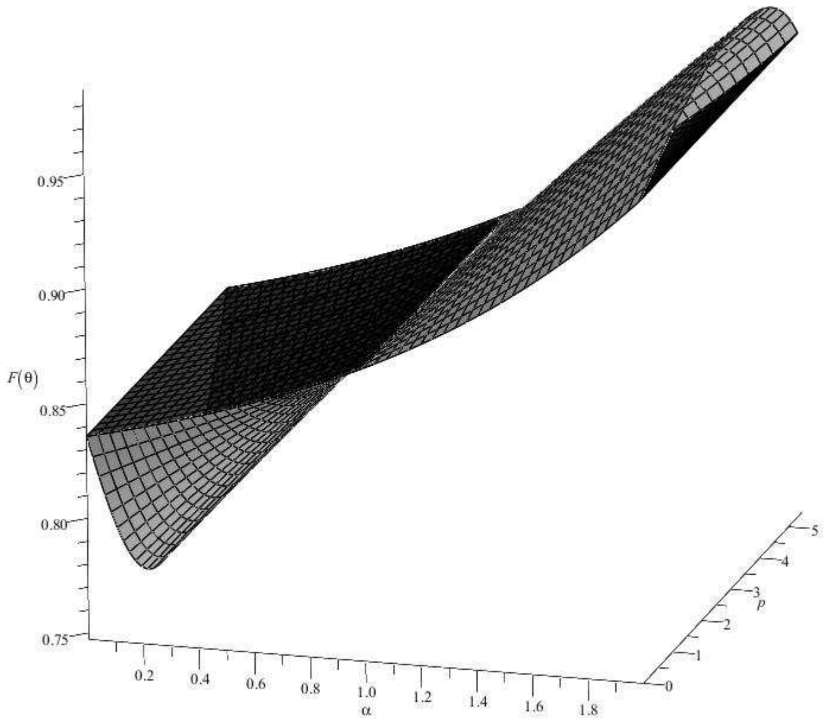

Figure 3.2.: Active CSPs under CSP tiering (gray) and NN (black) for different distributions of CSPs congestion sensitivity (α).
(The figure is derived for $\mu = 7/4$, $\lambda = 1$, $r = 1$, but qualitatively identical results are obtained for other parameter values.)

Assumption 2. *Content providers' congestion sensitivity, θ, is uniformly distributed such that $F(\theta) = \theta$.*

Under the present assumptions CSP tiering will lead to neither more nor less content variety. However, under a CSP tiering scenario the ISP can additionally extract rents from CSPs through sales of priority access. In the short run, it will do so by maximizing revenues from priority sales ($\Lambda\beta p$), which is achieved by

$$p = \left(1 - \sqrt{\frac{\bar{\theta}_Q}{\mu}}\right) r = \left(1 - \frac{w_{Q1}}{w_Q}\right) r \,. \tag{3.17}$$

Intuitively, this shows that the ISP can extract a fraction of the CSPs' gross advertisement revenue r. The proportion of the extracted rent depends on the congestion alleviation to the priority class compared to the average congestion level in the network.

Proposition 2 (ISP Preferred Regime)**.** *The ISP always prefers the CSP tiering regime because it can make extra profits by selling a priority transmission service to CSPs.*

Proof. Given the fact that $F(\bar{\theta}_N) = F(\bar{\theta}_Q)$ under the given assumptions, it follows that

$$\Pi_Q - \Pi_N = \Lambda\beta p = \mu r \left(1 + \frac{1}{1+\lambda} - \frac{2}{\sqrt{1+\lambda}}\right),$$

which is always greater than zero for $\mu, r, \lambda > 0$. □

3.3.2. Short Run Welfare Implications

This section deals with the short run effect of CSP tiering on welfare. Total welfare, W, is the sum of IUs' surplus, CSPs' surplus and the ISP's profit. To evaluate the transition from NN to a CSP tiering scenario one has to look at the difference in social surplus between the two regimes, which is given by

$$\Delta W = (U_Q - U_N) + (\Gamma_Q - \Gamma_N) + (\Pi_Q - \Pi_N)\ . \tag{3.18}$$

Recall that $U_Q = U_N = 0$, because IUs' surplus is always fully appropriated by the ISP. Conversely, this implies that any change in IUs' surplus is reflected in a change of the ISPs's profit.

Furthermore, $\Pi_Q - \Pi_N > 0$ according to Proposition 2, because the ISP makes extra revenues through selling priority access. What remains to be examined is the short run effect of CSP tiering on CSPs' surplus.

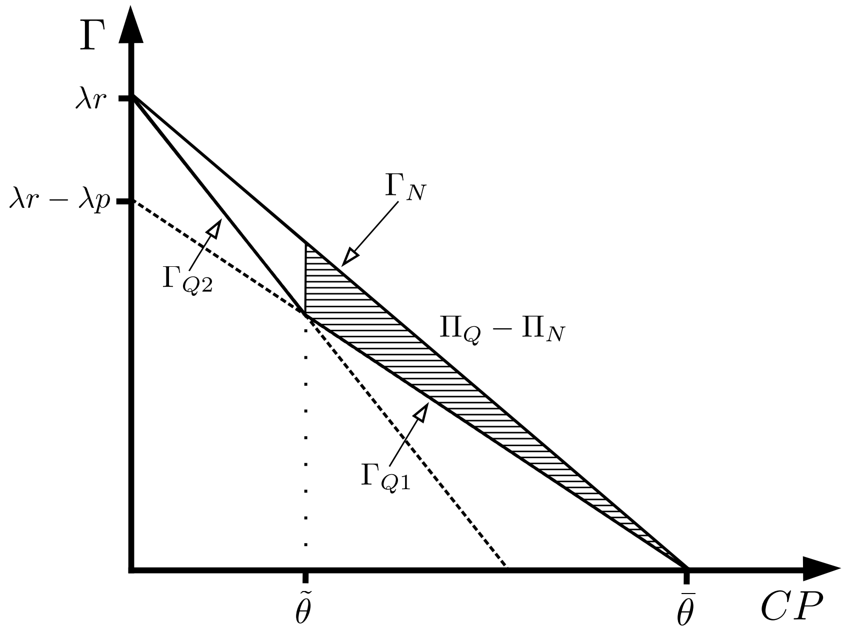

Figure 3.3.: The short run effect of CSP tiering on CSPs' surplus.

To this extent, consider Figure 3.3 and notice that those CSPs located at $\theta \in [0, \tilde{\theta})$ are evidently worse off under a CSP tiering regime, because for them network congestion has increased from w_N to w_{Q2}. This means a welfare loss, because this part of the surplus can not be expropriated by the ISP. Additionally, the CSPs' welfare loss increases with congestion sensitivity on the interval $\theta \in [0, \tilde{\theta})$. The business model of the provider located at $\theta = 0$ is not affected at all through congestion. Conversely, the business of the CSP at $\theta = \tilde{\theta}$ is suffering from congestion to such an extent that it is indifferent between staying in the BE class and buying priority access. However, notice that the welfare loss decreases for the CSPs in the priority class as $\theta \in [\tilde{\theta}, \bar{\theta})$ increases. To see this, recall from Proposition 1 that the last CSP to enter the market, $\bar{\theta}$, is identical under both regimes and receives a surplus of zero. For the last CSP that enters the market, the additional benefit through reduced congestion (compared to the NN regime) is equal to the price that it has to pay for priority access. Consequently, for all CSPs with less congestion sensitivity, $\theta \in [\tilde{\theta}, \bar{\theta})$, the price that is paid for priority is higher than the benefit of being in the priority class. However, for these providers the welfare loss is still smaller in the first priority class than in the BE class. In Figure 3.3 it can easily be seen, that

CSP $\tilde{\theta}$ incurs the greatest welfare loss. In summary, one can conclude that in the short run all active CSPs are (weakly) worse off under a CSP tiering regime.

However, recall that under the assumption of uniformly distributed congestion sensitivity no CSP is driven out of the market. The price that CSPs pay for priority access represents nothing more than a welfare shift to the ISP (hatched area in Figure 3.3). The sign of the overall welfare effect will therefore solely depend on the difference between the gross surplus gain through less congestion of those CSPs in the priority class and the gross surplus loss through increased congestion of those providers remaining in the BE class.

The following proposition summarizes the results w. r. t. Research Question 1.2.

Proposition 3 (Short run Welfare)**.** *If CSPs' congestion sensitivity is uniformly distributed, CSP tiering unambiguously increases welfare with respect to the NN regime in the short run, because congestion is alleviated for the most congestion sensitive CSPs in lieu of the less congestion sensitive CSPs. However, all CSPs are worse off under a CSP tiering regime because the increased surplus is expropriated by the ISP.*

Proof.

$$\begin{aligned}\Delta W &= (\Pi_Q - \Pi_N) + (\Gamma_Q - \Gamma_N) \\ &= \lambda r \left(\underbrace{(w_N - w_{Q1}) \int_{\tilde{\theta}}^{\bar{\theta}} \theta \, d\theta}_{\substack{\text{congestion alleviation} \\ \text{to priority class}}} \quad - \quad \underbrace{(w_{Q2} - w_N) \int_0^{\tilde{\theta}} \theta \, d\theta}_{\substack{\text{congestion aggravation} \\ \text{to best-effort class}}} \right) \qquad (3.19) \\ &= \frac{\lambda r}{2} \left(\left(\bar{\theta}^2 - \tilde{\theta}^2 \right) (w_N - w_{Q1}) - \tilde{\theta}^2 (w_{Q2} - w_N) \right).\end{aligned}$$

Thus,

$$\Delta W > 0 \Leftrightarrow \frac{w_N - w_{Q1}}{w_{Q2} - w_N} > \frac{\tilde{\theta}^2}{\bar{\theta}^2 - \tilde{\theta}^2}$$
$$\Leftrightarrow \frac{1-\beta}{\beta} > \frac{(1-\beta)^2}{1-(1-\beta)^2}$$
$$\Leftrightarrow 0 < \beta < 1\,.$$

Equation (3.19) reveals that the overall effect of CSP tiering on welfare depends on the relative size of the congestion alleviation effect to CSPs in the priority class (vertically hatched area in Figure 3.4) versus the congestion aggravation effect to CSPs in the BE class (horizontally hatched area in Figure 3.4). These effects relate directly to the main argument of proponents and opponents of NN, respectively. □

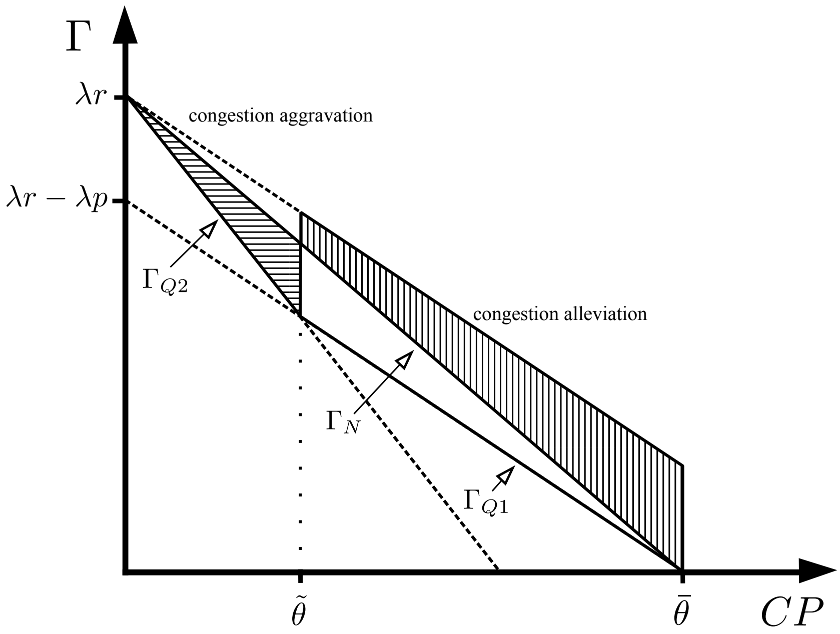

Figure 3.4.: Congestion alleviation vs. congestion aggravation effect of CSP tiering

Under a non-uniform distribution this welfare conclusion is not as clear-cut. If CSP tiering leads to more content variety, then those CSPs who are newly active in the market will enjoy a higher surplus under NN, and Proposition 3 is even strengthened. However, if there is a relatively large mass of congestion sensitive CSPs in the economy such that CSP tiering leads to less content variety, the associated welfare loss must be counterweighted with the welfare gain from better congestion allocation. In this case it is likely that Proposition 3 does not hold anymore.

3.4. Long Run Effects on Infrastructure Investments, Innovation and Welfare

In Section 1.3 it has been exemplified that ISPs are most concerned with the increasing costs of infrastructure investments. However, ISPs would also like to attract and stimulate entry of new (congestion sensitive) CSPs, because this is valued by customers and has a positive effect on the revenue stream from access fees. At the same time ISPs disapprove of CSPs who free-ride on their costly infrastructure. CSP tiering seems to be a viable way out of this dilemma, but it is unclear whether in the long run this regime will lead to more or less incentives for infrastructure investments compared to the NN regime.

This Section deals with long run investment incentives in network capacity (μ). Recall, that in the model, transmission capacity is represented by the average service rate at which customer requests can be handled by the network. An increase of μ allows the ISP to process more service requests to CSPs at a time, and consequently the average waiting time in the network (the proxy for congestion) will be reduced.

3.4.1. Investment Incentives

Formally, the ISP's investment decision is a discrete decision stage even before he announces any prices. The ISP first chooses the network capacity level, μ, and subsequently sets the customer access charge, a, and the priority price, p (under CSP tiering). In the subgame perfect equilibrium, the ISP will set the optimal capacity level, which is defined by the point where the marginal revenues of capacity expansion, $MR \equiv \partial\Pi/\partial\mu$, equal marginal costs, $MC \equiv \partial c(\mu)/\partial\mu$. Consequently, the ISP's optimal capacity level will be higher if marginal revenues from capacity expansion are higher. In both network regimes the following two marginal effects of capacity expansion on ISP revenue can be distinguished:

- The *variety incentive* $\left(v \cdot \partial F(\bar{\theta})/\partial\mu\right)$ denotes the ISP's marginal revenue effect on the customer access fee that results from the entry of new, congestion sensitive CSPs.
- The *congestion incentive* $(-\iota \cdot \partial w/\partial\mu)$ denotes the ISP's marginal revenue effect on the customer access fee that results from a change of the overall congestion level.

Furthermore, notice that under the assumption of a uniform distribution of CSPs' congestion sensitivity, these investment incentives are always positive and identical under both scenarios (i. e., NN and CSP tiering). Hence, potential differences in investments may only be a result of an additional investment incentive that an ISP has under CSP tiering:

- The *priority revenue incentive* $(\partial(\beta\Lambda p)/\partial\mu)$ denotes the ISP's marginal revenue effect from selling priority access.

Consequently, the sign of the priority revenue incentive determines the possible differences between investment incentives under CSP tiering and NN. The following proposition summarizes the results w. r. t. Research Question 2.1.

Proposition 4 (Investment Incentives)**.** *If the congestion sensitivity of CSPs is uniformly distributed, the ISP's optimal capacity level is higher under CSP tiering.*

Proof. Incentives to invest into network capacity are higher under CSP tiering iff marginal revenues from priority sales are greater than zero, provided that the ISP revenues are concave, and the costs of capacity expansion convex in μ. The latter is warranted by assumption. To ensure that the ISP's revenues are concave the property $\partial^2\Pi_Q/\partial\mu^2 \leq 0$ has to be fulfilled. The second-order condition is thus given by

$$\frac{\partial^2\Pi_Q}{\partial\mu^2} = -\frac{\iota(1+\lambda)}{\mu^3} + \underbrace{\left[\frac{\partial^2 r(\bar{\theta})}{\partial\bar{\theta}^2}\frac{\bar{\theta}}{2} + \frac{\partial r(\bar{\theta})}{\partial\bar{\theta}}\right]}_{A}\frac{\partial\bar{\theta}}{\partial\mu}\underbrace{\left(1+\frac{1}{(1+\lambda)} - \frac{2}{\sqrt{1+\lambda}}\right)}_{B} < 0.$$

Since $B \geq 0$ always holds, the ISP's revenues are concave if

$$\frac{\partial^2\Pi_Q}{\partial\mu^2} \leq 0 \begin{cases} A \leq 0 & \text{always} \\ A > 0 & \text{if} \quad \frac{\iota(1+\lambda)^2}{\mu^3 B} \geq A. \end{cases}$$

It is easy to see that $A \leq 0$ is warranted if ad revenues are decreasing (which is given by assumption) and concave (or not too convex). Otherwise we must assume, that the condition in the second case holds. However, alternatively it can also be assumed that the second-order condition holds locally around μ^*. Now consider $\Pi_Q - \Pi_N = \Lambda\beta p$. Differentiating with respect to μ yields

$$\frac{\partial(\Pi_Q - \Pi_N)}{\partial\mu} = \frac{\sqrt{\lambda+1}\left((\lambda+1) - \sqrt{\lambda+1}\right)}{(\lambda+1)\sqrt{\lambda+1}}\left[\frac{\partial r(\bar{\theta})}{\partial\bar{\theta}}\frac{\partial\bar{\theta}}{\partial\mu}\mu + r(\bar{\theta})\right].$$

The sign of the derivative is determined by the part in square brackets. Notice from (3.13) and (3.14) that $\partial\bar{\theta}/\partial\mu = 1/(\lambda+1)$ and $\mu = \bar{\theta}(\lambda+1)$. Consequently,

$$\frac{\partial(\Pi_Q - \Pi_N)}{\partial\mu} > 0 \Leftrightarrow \varepsilon^r = \frac{\partial r(\bar{\theta})}{\partial\bar{\theta}}\frac{\bar{\theta}}{r(\bar{\theta})} > -1\,.$$

Note that the gross industry advertisement revenue under NN is given by $R(\bar{\theta}) = \lambda r(F(\bar{\theta}))F(\bar{\theta})$. It is sensible to assume that $R(\cdot)$ does not decrease as more content becomes available. Thus, under the uniform distribution,

$$\frac{\partial R(\cdot)}{\partial\bar{\theta}} = \lambda\frac{\partial r(\cdot)}{\partial\bar{\theta}}\bar{\theta} + \lambda r(\cdot) > 0,$$

which holds iff $\varepsilon^r > -1$, in which case QoS tiering leads to more investments. □

This finding contrasts the results of Cheng et al. (2011) and Choi and Kim (2010). The reason is that this model explicitly accounts for the fact that more network capacity encourages the entry of new CSPs, whose additional demand keeps the value of the priority service high. As mentioned in Section 3.1, in Cheng et al. (2011) and Choi and Kim (2010) entry of new CSPs is not possible, and therefore it is more profitable to exploit the current CSP base and to keep network capacity scarce.

This picture may change under a non-uniform distribution function. In this case the mass of active CSPs may be different in the two scenarios, and consequently variety and congestion incentive may be not in line anymore. If the mass of congestion sensitive CSPs is very small, the priority revenue incentive can even be negative, because the ISP can sell priority access only to a very small number of customers. Hence, it seeks to make the priority class attractive to less congestion sensitive CSPs by keeping network capacity scarce and the congestion level high.

To follow this argumentation, consider Figure 3.5, which presents a numerical example of the marginal investment incentives for varying distributions of CSPs' congestion sensitivity. In line with Proposition 4, the variety and congestion incentive under either regime coincide under the uniform distribution of CSPs' congestion sensitivity ($\alpha = 1$), such that the positive priority revenue incentive is decisive for the higher investment incentives under CSP tiering.

Nevertheless, the more congestion sensitive CSPs are in the Internet economy ($\alpha < 1, \alpha \to 0$), the stronger is the variety incentive under NN compared to CSP tiering. On the other hand the priority revenue incentive, which is only present under CSP tiering, increases. As the variety incentive grows linearly in v and the priority revenue incentive grows linearly in r, NN may only lead to more infrastructure investments for $\alpha < 1$ if v is sufficiently larger than r. The intuition behind this result follows the classic two-sided logic (cf. Section 2.2). If the valuation from IUs for an additional CSP is high enough, it is profitable for the ISP to invest in additional capacity and stimulate entry of congestion sensitive CSPs on the other side of the market.

Conversely, when there are relatively less congestion sensitive CSPs in the economy ($\alpha > 1$), the variety and congestion incentive are slightly larger under CSP tiering while the priority incentive remains positive. In this case, CSP tiering provides unambiguously higher incentives for infrastructure investments.

However, when the mass of congestion sensitive CSPs becomes very small ($\alpha \gg 1$), the priority revenue incentive can indeed become negative, and eventually also the variety incentive under QoS tiering drops below the level under NN. In this case, it is likely that NN promotes investments in network infrastructure more.

In summary, one can conjecture that Proposition 4 holds locally around $\alpha = 1$ (Assumption 2), i.e., if the proportion of congestion sensitive CSPs to congestion insensitive CSPs is balanced.

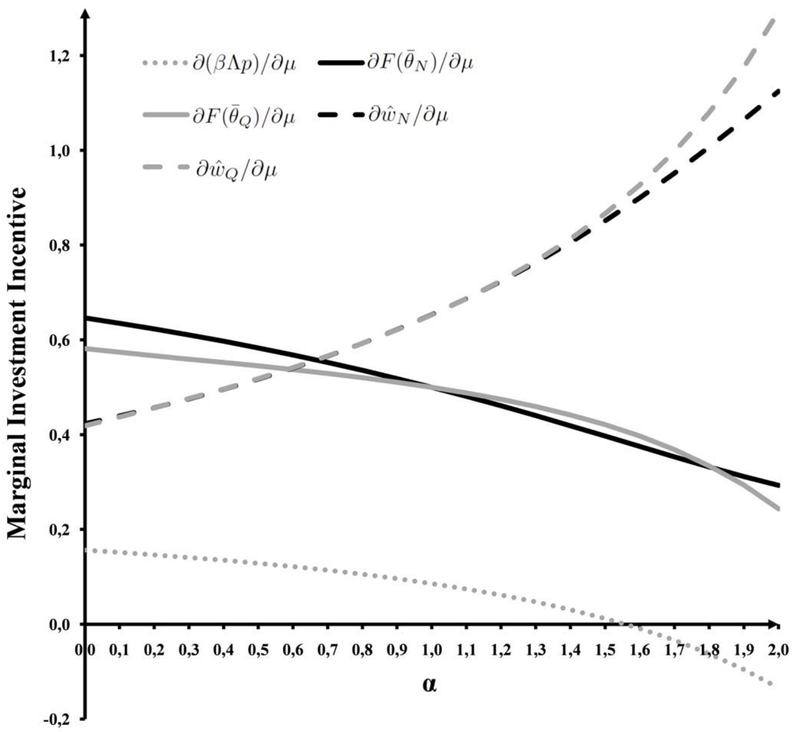

Figure 3.5.: ISP's marginal investment incentives under CSP tiering (gray) and NN (black) for different distributions of CSPs' congestion sensitivity (α).
(The figure is derived for $\mu = 7/4$, $\lambda = 1$, $r = 1$, $v = 1$.)

3.4.2. Innovation at the Edge and Long Run Welfare

The ISP's investments in network infrastructure have a direct effect on welfare. First, on the IU side of the market, customers experience lower network congestion (congestion incentive) and higher network effects (variety incentive). Second, on the CSP side of the market, the click-through rate increases due to lower network congestion, and consequently the slope of their surplus curve in both transmission classes increases, too. Nevertheless, all CSPs in the BE class and also some CSPs in the first priority class may still be worse off compared to NN. However, capacity expansion has an additional positive effect.
As a consequence of the overall decreased congestion level, both marginal CSPs, $\tilde{\theta}$ and $\bar{\theta}$, are shifted to the right. This means that new, highly congestion sensitive CSPs are able to enter the network. Figure 3.6 shows the effect of capacity expansion for CSPs under CSP tiering. In the context of the NN debate this has been referred to as 'innovation at the edge' (Jamison and Hauge, 2008). This finding illustrates the highly complementary character of network innovation (innovation at the core) and content and service innovation (innovation at the edge).

Obviously, the surplus of the new CSPs (crosswise hatched area), but also the surplus of some of the previously most congestion sensitive CSPs (vertically hatched area), are thus increased compared to a NN regime.

Accordingly, higher capacity levels will ceteris paribus lead to higher gross utility for consumers and CSPs, and are thus beneficial for welfare. The following proposition summarizes the findings w. r. t. Research Question 2.

Proposition 5 (Long Run Welfare)**.** *The regime that provides more incentives for infrastructure investments is more efficient in the long run. If the congestion sensitivity of CSPs is uniformly distributed, CSP tiering is more efficient and provides more content variety than NN.*

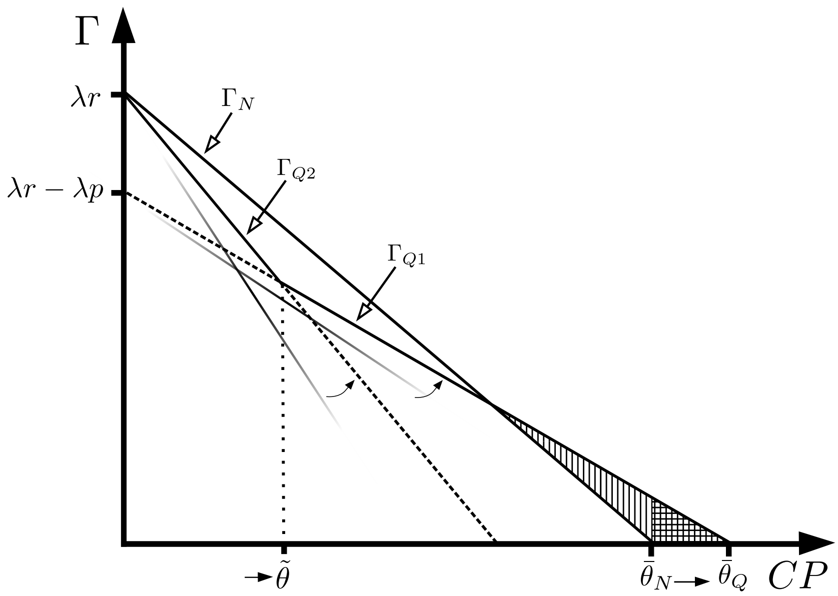

Figure 3.6.: The long run effect of CSP tiering on innovation and welfare.

Proof. To see that the overall congestion level, w, decreases with capacity expansion, one has to show that $\partial w/\partial \mu = \partial(1/(\mu-\Lambda))/\partial\mu < 0$.

Notice that $\Lambda = \bar{\theta}\lambda = \lambda\mu/\lambda+1$, so that $\partial\Lambda/\partial\mu = \lambda/(\lambda+1) < 1$. Therefore, it holds that $\partial(\mu-\Lambda)/\partial\mu > 0$ and consequently, $\partial(1/(\mu-\Lambda))/\partial\mu < 0$. By equation (3.2) and (3.3) it is easy to see that the gross utility of customers and CSPs increases as the congestion level decreases. The homogeneity of customers allows the ISP to fully expropriate the additional customer utility. Capacity expansion also increases the amount of active CSPs, since $\partial\bar{\theta}/\partial\mu > 0$ by equation (3.14). Before the capacity expansion occurred, these CSPs had a surplus of zero and are therefore unambiguously better off.

If QoS tiering provides a higher capacity level ($\mu_Q^* > \mu_N^*$), the critical CSP that is just equally well off as under NN is determined by the equation $(1-\breve{\theta}w_N)\lambda r_N = (1-\breve{\theta}w_{Q1})\lambda r_Q - \lambda p$. Inserting (3.17) and reformulating yields:

$$\breve{\theta} = \frac{(r_N - r_Q)\,w_Q + (w_Q - w_{Q1})}{w_Q\,(w_N r_N - w_Q r_Q)}. \tag{3.20}$$

Because $\mu_Q^* > \mu_N^*$, it follows directly that $w_N > w_Q$, $\bar{\theta}_N < \bar{\theta}_Q$ and thus $r_N \geq r_Q$. It is easy to see that $0 < \breve{\theta} < \bar{\theta}_Q = 1/w_Q$. Therefore, all CSPs in the interval $(\breve{\theta}, \bar{\theta}_Q]$ are better off than under NN. □

CSP tiering is consequently the more efficient regime in the long run and, by Proposition 3, also in the short run if CSPs' congestion sensitivity is uniformly distributed. Nevertheless, all CSPs in the interval $(0, \breve{\theta})$ are still worse off compared to the NN regime.

3.5. Alternative Model Variants

3.5.1. Model with Competitive Clicks

As explained in Section 3.2 this chapter is not intended to study the effect of CSP tiering on the direct competition between otherwise similar CSPs. Nevertheless it has to be controlled for the alternate assumption, that the consumers' total number of clicks is fixed and thus λ diminishes as the variety of content and services in the Internet economy increases. This allows to analyze a general notion of competitive pressure among the CSPs.

To analyze this scenario assume each IU spends an exogenous amount of clicks Λ on the Internet, which he evenly distributes among the available CSPs. Therefore, the arrival rate at each CSP is no $\lambda = \Lambda/F(\bar{\theta})$. Observe now that λ decreases as the number of active CSPs increases. However, this alternative model variant is still in line with Assumption 1, because each active CSP receives the same number of clicks from each customer. In conjunction with Assumption 2 this model variant yields the following results:

$$F(\bar{\theta}_N) = \bar{\theta}_N = \mu - \Lambda \tag{3.21}$$

$$F(\bar{\theta}_Q) = \bar{\theta}_Q = \mu - \Lambda \tag{3.22}$$

$$F(\tilde{\theta}) = \tilde{\theta} = \frac{p}{r-p}\frac{\mu-\Lambda}{\Lambda}\bar{\theta}_Q\,. \tag{3.23}$$

Obviously, it holds that $\bar{\theta}_N = \bar{\theta}_Q = 1/w_N$, which is exactly in line with Proposition 1 of the base model.

Furthermore, the optimal priority price, which maximizes $\Lambda\beta p$, is given by

$$p = \frac{\mu - \sqrt{\mu(\mu-1)}}{\mu} r = \left(1 - \sqrt{\frac{\bar{\theta}_Q}{\mu}}\right) r = \left(1 - \frac{w_{Q1}}{\hat{w}_Q}\right) r, \tag{3.24}$$

which is exactly the price structure that is described by (3.17). Consequently this model variant must yield the same qualitative results.

3.5.2. Model with Congestion Sensitive Consumers

In the base model network congestion is incorporated in the IUs utility function in the form of the average waiting time in the network. As explained in Section 3.2 this means that IUs dislike congestion independently from the CSP's innate sensitivity towards congestion. The following extension of the base model discusses the case of congestion sensitive IUs. More precisely, customers perceive network congestion more bothersome if the congestion sensitivity of the respective CSP is high. Note that consumers are still homogeneous w. r. t. congestion sensitivity. Now, however, congestion does not only reduce their click-through-rate on the CSP side but also their own valuation of each CSP.

To capture this effect, consider the IUs' utility function from (3.3), but now assume in contrast to the base model that consumers, instead of the average congestion level, evaluate congestion by $\hat{w}_N = \int_{\theta=0}^{\bar{\theta}} w_N \theta f(\theta) d\theta$ and $\hat{w}_Q = \int_{\theta=0}^{\tilde{\theta}} w_{Q2} \theta f(\theta) d\theta + \int_{\theta=\tilde{\theta}}^{\bar{\theta}} w_{Q1} \theta f(\theta) d\theta$, respectively.

It follows that $\hat{w}_N \geq \hat{w}_Q$ for $\bar{\theta}_Q = \bar{\theta}_N = \bar{\theta}$ and $\mu_N = \mu_Q = \mu$:

Proof.

$$\hat{w}_N \geq \hat{w}_Q \Leftrightarrow \frac{w_N - w_{Q1}}{w_{Q2} - w_N} \geq \frac{\int_{\theta=0}^{\tilde{\theta}} \theta f(\theta) d\theta}{\int_{\theta=\tilde{\theta}}^{\bar{\theta}} \theta f(\theta) d\theta}$$

Substituting $w_N = \beta w_{Q1} + (1-\beta) w_{Q2}$ and $\beta = 1 - F(\tilde{\theta})/F(\bar{\theta})$, and integrating the right hand side yields:

$$\frac{F(\tilde{\theta})}{F(\bar{\theta}) - F(\tilde{\theta})} \geq \frac{\tilde{\theta} F(\tilde{\theta}) - \int_{\theta=0}^{\tilde{\theta}} F(\theta) d\theta}{\bar{\theta} F(\bar{\theta}) - \tilde{\theta} F(\tilde{\theta}) - \int_{\theta=\tilde{\theta}}^{\bar{\theta}} F(\theta) d\theta}$$
$$\Leftrightarrow F(\bar{\theta})(\bar{\theta} - \tilde{\theta}) + \frac{F(\bar{\theta}) - F(\tilde{\theta})}{F(\tilde{\theta})} \int_{\theta=0}^{\tilde{\theta}} F(\theta) d\theta \geq \int_{\theta=\tilde{\theta}}^{\bar{\theta}} F(\theta) d\theta$$

From $F(\bar{\theta})(\bar{\theta} - \tilde{\theta}) \geq \int_{\theta=\tilde{\theta}}^{\bar{\theta}} F(\theta) d\theta$ and from $\bar{\theta} \geq \tilde{\theta}$ as well as the monotonicity of the distribution function F, it follows that the inequality is always satisfied. □

Consequently, if consumers are aware of the congestion sensitivity of content and services, CSP tiering yields a higher IU utility than NN. This alternative model variant introduces an additional short run welfare gain. Compared to the base model CSP tiering is even more likely to be welfare superior compared to NN, because the congestion alleviation becomes even stronger. However, recall that the ISP still extracts the whole consumer surplus due to the assumption of IU homogeneity.

Therefore, as mentioned in Section 3.2, the base model has to be considered as the more conservative model variant with respect to the analysis of the potential benefits of CSP tiering.

3.6. Minimum Quality Standards

In the context of CSP tiering with $M/M/1$ queuing, priority price regulation is not a suitable policy instrument, because in the short run social and private incentives are in line. To see this, note that the social planner seeks to set the regulated priority price such that the socially optimal share of CSPs selects the priority transmission class. Thereby CSPs' gross surplus is maximized.
However, the ISP pursues the same goal and consequently already chooses the optimal price himself, because it can extract a fraction of the CSPs' surplus. In contrast to the optimal level of the priority price, the social and private level of infrastructure investments are not in line.

Proposition 6 (Efficient Investments)**.** *The social planner has a higher incentive to invest in network capacity than the ISP. This result holds for both network regimes, QoS tiering and NN.*

Proof. Each regime is considered separately, and it is exemplified that the conditions with respect to efficient investments coincide. First, the conditions for which $\partial(W_N - \Pi_N)/\partial\mu$ is larger than zero are derived:

$$
\begin{aligned}
W_N - \Pi_N &= \frac{\lambda}{2(\lambda+1)}\mu r(\bar{\theta}) \\
\frac{\partial\left(W_N - \Pi_N\right)}{\partial\mu} &> 0 \\
\Leftrightarrow \quad \frac{\partial r(\bar{\theta})}{\partial\bar{\theta}}\frac{\bar{\theta}}{r(\bar{\theta})} &> -1 \\
\Leftrightarrow \quad \varepsilon^r &> -1
\end{aligned}
$$

The difference of private and efficient investment incentives under the CSP tiering regime is:

$$
\begin{aligned}
W_Q - \Pi_Q =& \frac{\sqrt{\lambda+1}-1}{\lambda+1}\mu r(\bar{\theta}) \\
\frac{\partial\left(W_Q - \Pi_Q\right)}{\partial\mu} =& \frac{\sqrt{\lambda+1}-1}{\lambda+1}\left(\frac{\partial r(\bar{\theta})}{\partial\bar{\theta}}\mu + r(\bar{\theta})\right) > 0 \\
\Leftrightarrow& \frac{\partial r(\bar{\theta})}{\partial\bar{\theta}}\frac{\bar{\theta}}{r(\bar{\theta})} > -1 \\
\Leftrightarrow& \varepsilon^r > -1
\end{aligned}
$$

By the same argument as in the proof of Proposition 4, only $\varepsilon^r > -1$ is feasible and thus the proposition obtains. □

Consequently, w. r. t. the ISP's investments in infrastructure, there is amended concern on the side of CSP tiering opponents. Section 2.3.1 introduced the metaphor of the dirt road fallacy and Section 2.4 showed that the biggest concern under QoS tiering in general is related to ISPs keeping capacity scarce (underprovisioning). As a regulatory remedy, politicians could opt for a MQS of the BE class (Brennan, 2011). Under the new European legislative framework (cf. Section 2.4) a MQS policy is already feasible. After all, MQSs have been found to be welfare-enhancing in competitive settings (Ronnen, 1991).

Some politicians go even one step further and demand that new QoS technologies should never lead to a lower quality than under NN. They argue that ISPs should be allowed to earn additional revenues by selling a higher quality, but at the same time it should be forbidden to reduce the BE quality below the level of the status quo (NN). Consequently, under the CSP tiering regime no CSP would be set at a disadvantage anymore. In the model this would mean that the MQS would be set in a way so that under CSP tiering the ISP is required to offer CSPs a congestion level in the BE class that is at least as low as the equilibrium BE congestion level under NN.

At the same time, opponents of NN regulation have often objected that NN forces ISPs to invest above the efficient level, which is known as overprovisioning. Forcing them to keep up this high level of investments under CSP tiering can therefore not be efficient from the point of view of the ISP.

In order to meet this MQS, the ISP is required to increase the network's capacity, potentially to the extent that the gap between the level of private and efficient investments is closed. More precisely, by requiring the MQS $w_N(\mu_N^*) \equiv w_{Q2}(\mu_{MQS})$ the regulator implicitly defines the new capacity level $\mu_{MQS} > \mu_N^*$.[5]

By Propositions 4 and 6 the order of relevant capacity levels is

$$\mu_Q^{**} > \mu_Q^* > \mu_N^*.$$

Recall that $\mu_{MQS} > \mu_N^*$, and thus one can differentiate between three different cases.

1. If $\mu_Q^* \geq \mu_{MQS}$, the MQS is not a binding condition for the ISP's investment decision and hence is simply ineffective.

2. If $\mu_Q^{**} \geq \mu_{MQS} > \mu_Q^*$, the MQS is effective in raising the ISP's network capacity level, potentially up to the efficient level.

3. If $\mu_{MQS} > \mu_Q^{**}$, the MQS policy may lead to an excessive investment in network infrastructure.

In summary, MQSs are only effective in one out of three cases. This result shows that MQS can also lead to an inefficient outcome. Therefore the level of the MQS has to be careful calibrated and it is questionable whether the congestion level under the status quo (NN) is an appropriate objective for the BE class under CSP tiering.
However, even if a regulatory agency would be able to determine the efficient level of capacity

[5] One asterisk denotes the equilibrium capacity level, whereas two asterisks denote the socially optimal capacity level.

(μ_Q^{**}), MQS are only an indirect tool to achieve the desired capacity level. The following proposition summarizes the findings w. r. t. Research Question 3.

Proposition 7 (Minimum Quality Standard Regulation)**.** *An MQS policy which requires the ISP to guarantee a BE congestion level under CSP tiering which is equal to the equilibrium congestion level under NN may increase welfare, but it may also lead to excessive investments or be ineffective.*

Proof. To show that the ISP under a MQS enforcement of $w_{Q2} = w_N$ has a higher incentive to invest in capacity than under NN one has to show that $\mu_{MQS} > \mu_N^*$.

$$w_{Q2} = w_N$$
$$\Leftrightarrow \quad \frac{\mu_{MQS}}{\mu_{MQS} - \lambda\bar{\theta}} \frac{1}{\mu_{MQS} - \lambda\beta\bar{\theta}} = \frac{1}{\mu_N^* - \lambda\bar{\theta}_N}$$
$$\Leftrightarrow \quad \mu_{MQS} = \frac{1+\lambda}{1+\lambda(1-\beta)}\mu_N^*$$

Since $\beta < 1$, it is easy to see, that $\mu_{MQS} > \mu_N^*$ always holds true. □

3.7. Implications

This chapter contributes to the debate on NN by providing a formal framework that extends the insights gained from the existing theoretical models in several ways. The analysis focuses on the relationship between CSPs and a monopolistic ISP, and compares NN to a CSP tiering regime in which CSPs may pay for the prioritized transmission of their data packets on a non-discriminatory basis. The model adds to the field by explicitly considering the effect of CSP tiering on content and service variety and by taking into account that CSPs are heterogeneous w. r. t. their sensitivity towards congestion. Due to the use of basic insights from queuing theory, the model explicitly takes into account the negative externality that prioritization has on the remaining BE class.

The results of the comparison between the two network regimes depends on the distribution of CSPs' congestion sensitivity in the Internet economy. In particular, under the neutral reference case where CSPs' congestion sensitivity is uniformly distributed, CSP tiering increases welfare in the short run because the installed level of network capacity is used more efficiently: Network congestion is re-allocated, such that it is alleviated for the most congestion sensitive CSPs. This offsets the congestion aggravation for the CSPs in the remaining BE class. However, CSP tiering does not immediately promote the entry of new content providers with innovative services that are even more congestion sensitive. In fact, in the short run, all CSPs are likely to be worse off under a CSP tiering regime because the ISP is able to expropriate some of the CSPs' surplus through priority pricing. Consequently, the ISP always prefers the CSP tiering regime. It is subject to the authority of policy makers to evaluate the shift of surplus from CSPs to ISPs, which is welfare neutral per se, but lies at the heart of the NN debate.

On the other hand, ISPs argue that they will use the additional revenues to raise investments in broadband infrastructure. This is true for the reference case of uniformly distributed congestion sensitivities, but may not hold for more skewed distribution functions.

In sum, CSP tiering is likely to be the more efficient regime if the proportion of congestion sensitive to congestion insensitive CSPs is balanced. In this case, the ISP invests more in broadband infrastructure, and thereby allows for entry of new, congestion sensitive CSPs in the short run. Therefore, the highly congestion sensitive CSPs in particular will be better off under CSP tiering, and hence it is not surprising that Google and Verizon have privately agreed on a tiered system (Wyatt, 2010). According to Definition 1.4, Google CEO Eric Schmidt argues that such an agreement would be in line with NN, because it does not discriminate against specific CSPs.

Furthermore, the analysis reveals that the level of private investments is generally not efficient. A MQS policy that requires the ISP to guarantee a congestion level in the BE class under

CSP tiering which is at least as low as the BE congestion level under NN is not sufficient to guarantee efficient infrastructure investments.

In conclusion, while the model results show that some of the objections to CSP tiering are justified, a strong case for a tiered network also exists. The potential dangers of a CSP tiering regime can be overcome by transparency obligations or MQSs. Under competition ISPs will try to attract customers by offering them more content variety and a lower average congestion level than their competitor. This will boost their investment incentives. Likewise, the ISPs will also lower the customers' access charge, and therefore some of the ISPs' rent is shifted towards the consumers. However, competition between ISPs does not change the main insights of the analysis under monopoly if one makes the reasonable assumption that CSPs multihome (i.e., are connected with BE to every ISP) whereas consumers singlehome (i.e., are connected to one ISP exclusively). In this case, every CSP would again face a terminating monopoly over the connected consumers at each ISP, leaving the previously described relationship between the ISP and the CSPs intact. Consequently, the result that QoS tiering is the ISPs' preferred regime and that it will lead to more investment and content variety due to the additional priority revenue incentive remains unchanged. Hence, there is no reason to believe that competition between ISPs will warrant NN. In reverse, the prohibition of QoS tiering (pay-for priority), which has been proposed by the FCC for fixed line networks, can eventually be harmful to content variety, broadband investment and welfare.

Chapter 4.

Internet User Tiering

“The Internet is increasingly becoming the dominant medium binding us. The neutral communications medium is essential to our society. It is the basis of a fair competitive market economy. It is the basis of democracy, by which a community should decide what to do. It is the basis of science, by which humankind should decide what is true.”

Sir Tim Berners Lee (Lee, 2006)

THIS chapter investigates the factors that may influence a transition to a differentiated service regime in the Internet in a scenario where IUs pay for preferential treatment.

4.1. Motivation

In the status quo of the Internet pricing of IUs is a simple access fee model. They pay for bare access to the (managed) BE network, but there are no extra fees for preferential treatment.

One possible deviation from the status quo discussed in Section 2 is called IU tiering. The introduction of pricing schemes that are based on QoS technology are not restricted to the CSP side of the market. Likewise, some definitions of NN (cf. Definitions 1.1 and 1.2) would identify the introduction of QoS based Internet tariffs for consumers as a violation of the NN principle. However, Definition 1.4 explicitly includes QoS under the umbrella of NN. Also, regulators in Europe are less concerned about this specific scenario of NNN. This has two reasons: First, consumers can decide whether they wish to be prioritized or not. Second, many of the threats identified in Section 2.4 are irrelevant, because IU tiering is a one-sided pricing scheme (i. e., only IU pay). Therefore, regulators are less concerned about reductions in innovation at the edge, because all CSPs face a level playing field.

Recall from Section 2.3.2 that the reason for a potential deviation from the status quo is based on the claim of ISPs that the current pricing regime for Internet access cannot cover the costs of future infrastructure investments. As a consequence consumers would face a congestion problem, because ISPs will not be able anymore to ensure the current quality of Internet access services by overprovisioning of network capacity.

4.1.1. Quality Provision Mechanism

In many countries Internet tariffs are predominantly based on flat rates. ISPs discriminate solely via the maximum bandwidth of the physical connection. There is currently no widespread introduction of QoS based Internet access tariffs in the market. The example of the British ISP Plusnet (cf. Section 2.3.2), however, shows that an IU tiering scenario, in contrast to a CSP tiering scenario, is not far fetched and has already been brought to the market.

Obviously, the general logic of queuing models also applies to a scenario in which some users get preferential treatment in the network and others do not. A prioritization of some users inevitably has a negative effect on the remaining BE users. Note that this effect is not necessarily related to the bandwidth of an IU connection. Consider for example an IU using a narrow-band Integrated Services Digital Network (ISDN) connection and, for example, an IU using a broadband Very High Data Rate Digital Subscriber Line (VDSL) connection. Both users participate in an online auction for a single item. Before the auction ends they want to send a last second bid ('sniping') to the online auction platform. Assume further, that the narrow-band IU bought preferential treatment (priority access) for a premium fee from his ISP while the broadband IU did not. In this special situation bandwidth is not the crucial factor for the bid to arrive in time, because the transmission of a 'single click' is not constrained by the ISDN connection bandwidth. However, the priority transmission of the narrow-band IU yields that his bid arrives at the online auction platform in time, whereas the broadband IU looses the auction because his bid arrives after the auction is finished.

With this example in mind, it becomes obvious that 'quality' of an Internet connection is in the eye of the beholder. While mobile users may value a high coverage of the network and a stable connection on the move (reliability), online brokers and online gamers may focus on the aspect of latency, whereas file sharers may base their decision on the bandwidth of the available Internet access alternatives.

A similar logic holds true for the availability of specific services and content. Business customers may value the availability of high quality video telephony and Virtual Private Network (VPN) connections, while private customers value access to big social networks (e. g., Facebook) and real time entertainment. Consumers are undoubtedly heterogeneous w. r. t. their preferences about Internet access in many aspects. However, in the BE Internet the network technology and the available tariffs for Internet access currently cannot account for such differences. QoS could be a viable way to account for individual preferences of IUs and to monetize them. Some IUs may also have a constrained budget and want to buy cheaper Internet access. Current Internet pricing schemes offer only 'full access'. The cite from Nelly Kroes in Section 2.3.2 shows that she is concerned that NN regulation may hinder new pricing schemes that allow users to buy access to only a sub-set of available content and services, or unprioritized access to the Internet for a lower price.

The economic literature about QoS tiering can be divided in two distinct groups. One strand of the literature (e. g., Bandyopadhyay and Cheng, 2006; Cheng et al., 2011; Choi and Kim, 2010; Krämer and Wiewiorra, 2012; Reggiani and Valletti, 2012) employs a relative quality provision mechanism ($M/M/1$) to fix the relationship between priority and BE transmission, whereas the other strand (e. g., Economides and Tag, 2012; Njoroge et al., 2010) employs a dedicated quality provision mechanism. With dedicated quality provision, the adverse effect of priority provision on the BE quality is constrained. One could assume that the ISP guarantees a certain level of quality (Service Level Agreement (SLA)), or assigns a fixed share of the overall capacity explicitly to priority and BE users respectively (Paris Metro pricing). While the $M/M/1$ model relies on adding priority information to data packets, such that they can be processed ahead of other data packets, Paris Metro pricing relies on the split of network capacity between multiple service classes. The price for each service class implicitly regulates how congested the respective class is. In this respect Paris Metro pricing can be thought of multiple BE effort networks owned by the same ISP, which achieves quality differentiation solely through price differentiation.

Besides technical differences (e. g., costs of providing QoS with either mechanism) it is not clear ex-ante how users perceive the two classes of quality provision mechanisms. Therefore, it is important to explore the acceptance of both mechanisms in direct relation to NN. It is also possible that the WTP depends on the quality provision mechanism, too. The implementation of a quality provision mechanism that is preferred by IU could help to facilitate a successful transition to an IU tiering scenario.

Research Question 4. *Are relative and dedicated quality mechanisms perceived differently by consumers, and if so why?*

Research Question 5. *Is it likely that ISPs could increase their profits by offering IU tiering?*

4.1.2. Fairness

It has been exemplified that consumers can react very negatively to changes in pricing. In Section 2.3.2 the example of big Canadian ISPs was mentioned, which tried to revert the flat rate dominated Internet access market back to a metered pricing scheme. They faced tremendous public protest, and the regulator finally stopped the ISPs from doing so.
The topic of NN became a politically heavily debated matter, and this example demonstrates that changes in the network level (reduced data caps, different wholesale pricing) can result in unforeseen reactions of consumers. IU tiering could trigger similar reactions if one does not understand how users perceive prioritization and possible pricing schemes that come along with such a change in network regime.

For these reasons it is not very surprising that ISPs in general are very reluctant to talk about this scenario publicly and seem to primarily focus their lobbying efforts on the introduction of QoS on the CSP side of the market.[1] As illustrated before this may be due to different reasons. The most striking argument, however, appears to be the perception of fairness w. r. t. the distribution of network resources and the provision of quality.

The following overview of the most relevant fairness concepts and the perception of price fairness is based on the papers by Xia et al. (2004) and Martín-Ruiz and Rondán-Cataluña (2008).
In general, consumers are more likely to perceive price increases as unfair if they are justified with scarcity of the good or are triggered by higher demand (Frey and Pommerehne, 1993). If the price of soda cans is raised at a very hot day because the seller expects higher demand, consumers perceive this price increase as unfair. However, Vaidyanathan and Aggarwal (2003) find that consumers are less likely to perceive price increases as unfair, if the reason for the increase is out of the seller's control. (e. g. , the price peak of computer hard disk drives after

[1] However, mobile telecommunications providers already offer tariff options that could be considered a deviation from the classic IU pricing and a violation of NN. T-mobile Germany sells a VoIP option that allows customers to use VoIP applications (which is excluded by the general terms of service). However, this can be considered as anti-competitive behavior and is therefore not an inherent concern under IU tiering.

a flooding of the production facilities). Bolton et al. (2003) find that cosumers are more likely to accept price differences that are based on differences in the quality of the products.
With respect to Internet pricing both arguments may be applicable. Due to excessive demand the good is scarce, but on the other hand the introduction of QoS would introduce quality differentiation to the otherwise homogeneous product.

Consumers generally use different reference points to evaluate prices of products and services. They rely on past prices and the prices of competitors (Bolton et al., 2003). Furthermore, they compare their present outcome with their past outcome, but they also compare their personal present outcome with the present outcome of other consumers (Xia et al., 2004).
For a study of IU tiering it is therefore very important to provide participants with a neutral and equal reference point.

Services (like Internet access) in contrast to durable goods are valued differently by consumers. Costs of durable goods seem to be more acceptable and therefore are less likely to be perceived as unfair. “Specifically, goods have salient material costs that serve as reference points for selling prices and price increases and that may distract consumers from other costs” (Bolton and Alba, 2006, p.264).
Since consumers may underestimate the ISPs’ costs for the the necessary infrastructure, Internet access services would have an inherent disadvantage in that respect.

The literature on fairness, in particular price fairness, offers a rich set of different theories and fairness concepts. Xia et al., p.1 summarize the ‘classic’ concept of fairness “as a judgment of whether an outcome and /or process to reach an outcome are reasonable, acceptable, or just”. Moreover, Xia et al. state that unfairness is not necessarily the other side of the coin. While it is often clear to customers what they perceive as unfair it seems to be difficult for them to define what is fair. They argue that the strong emotional reaction to unfair situations is the main reason why consumers can judge such situations more easily. Obviously humans are influenced by different aspects when they evaluate whether an outcome is perceived as fair.

Thus, it is helpful to understand which aspects may influence their perception of differentiated Internet access.

This thesis relies on the theoretical framework of justice theory to capture what consumers perceive as 'just' or 'fair' in relation to Internet pricing. The concept of justice can be differentiated into distributive justice (Adams, 1965), procedural justice (Homans, 1961) and interactional justice.
Distributive justice is related to the outcome of allocations. The 'Equity' principle suggests that "[...]a customer's reward should equal his or her contribution to the exchange"(Martín-Ruiz and Rondán-Cataluña, 2008, p.328), whereas the 'Equality' principle suggests that "[...]different customers are entitled to receive the same outcome"(Martín-Ruiz and Rondán-Cataluña, 2008, p.328). Procedural justice is related to the system that is used to determine the outcome of an allocation process, whereas the concept of interactional justice is related to "the interpersonal treatment individuals receive from service providers" (Martín-Ruiz and Rondán-Cataluña, 2008, p.328).

A better understanding of the different aspects of fairness that influence the acceptance of a certain quality provision mechanism and the WTP for priority access could help ISPs to adress IUs concerns more accurately ex-ante (e. g., via advertisements) and to plan their network systems' design.

Research Question 6. *Which aspects of fairness influence IUs' preference for IU tiering?*

Research Question 7. *Which aspects of fairness influence IUs' WTP for priority and the necessary monetary compensation for BE under IU tiering?*

4.1.3. Combining Theories

In an effort to understand how IUs would perceive a transition from a BE Internet access to an IU tiering scenario, and how they evaluate NN in general, 'fairness' is only one piece of the

puzzle. Research on QoS perception (e. g., Bouch and Sasse, 1999; Bouch et al., 2000) shows that additional dimensions like 'task importance', 'feedback' and the consumers' budget are also important factors. Sometimes this field of research yields qualitatively similar results to the economic theory of tariff biases. Lambrecht and Skiera (2006a) find that predictability of the expenditure (insurance effect) and the avoidance of continued costs signals (taximeter effect) are important in Internet tariff decisions of customers.
Moreover, any NNN scenario is currently hypothetical and thus any research project w. r. t. user perception can only rely on data based on hypothetical scenarios. The line of research related to a similar kind of questions is called 'technology acceptance theory' (e. g., Venkatesh et al., 2003). In contrast to this study, Technology Acceptance Models (TAMs) consider the usage intention of new Information Technology (IT) systems in an business environment. Therefore many aspects are not applicable to a purely hypothetical service (e. g., social influence[2]).

Obviously, none of these theories covers all relevant aspects of NN or IU tiering. On the other hand, considering all factors of each theory, however, would overload the analysis by irrelevant factors w. r. t. the NN debate.

The following sections present a modular model that was created in an effort to combine the relevant factors from the theoretical approaches shown in Figure 4.1.

4.2. Methodology

Psychological concepts like fairness are not part of traditional economic theory. To assess the economic impact of such latent psychological factors, this thesis employs a so-called 'quasi experimental' approach (e. g., Lambrecht and Skiera, 2006b). Quasi experiments are an empirical methodology often used in form of a survey with hypothetical choice scenarios.

[2] If a technology has not been launched on the market yet, there cannot be any impression based on experience about it among social relevant peers that could influence the consumers opinion.

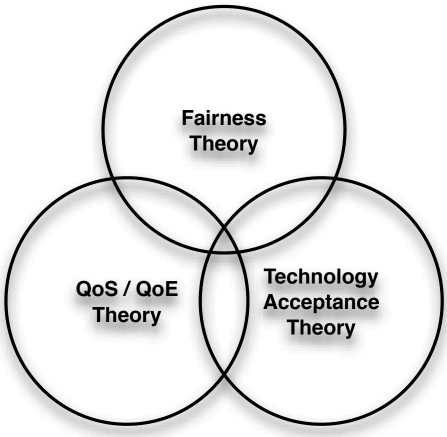

Figure 4.1.: Combining different theories

Thereby, respondents have to indicate their own expected decision in the presented hypothetical situations. Along with the choice scenarios, additional explanatory variables are surveyed that are assumed to have an influence on the decision at hand. This methodology is considered to be only a ‘quasi’ experiment because it lacks the control of laboratory experiments. In contrast to a laboratory environment, the researcher neither observes the conditions under which the respondent answers the questions, nor controls possible interaction of the respondents with others, and he lacks full control over a possible treatment assignment. However, in combination with a web-based survey design, this approach allows to gather large sample data using a standardized interface (in contrast to personal interviews or focus group studies). The data acquisition procedure for this thesis was conducted via a online panel among 1035 representative IUs, of which all are responsible for their own Internet access service.

The measurement model and the statistical methodology that is suitable to answer the aforementioned research questions must be able to simultaneously examine the influence of unobservable (latent) and observable parameters on the network regime decision of IU as well as on their WTP for a higher level of service. Latent variables, like fairness, can only be reliably measured by multi-item scales, which cannot be sophisticatedly incorporated into standard regression analysis. An *item* is a single verbal statement that is assumed to be an indicator of the unobservable variable (e. g. , fairness). Respondents have to indicate their level of agreement to each individual item (statement) on a five point Likert scale Likert (1932). Multi-item scales are used because single items are prone to measurement errors, because respondents may differ in their evaluation of a single aspect of an unobservable variable, but not in their evaluation of the unobservable variable as a whole. The set of items that is assumed to measure a latent variable is called a *construct*.

Structural Equation Models (SEMs) are well suited for this kind of mixed hypothesis testing. SEMs are a multivariate methodology to represent and test a theoretical network of relations between observed and unobserved variables (e. g., Bollen, 1989; MacCallum and Austin, 2000). Hoyle (2011) gives an introduction into SEM that can be briefly summarized as follows: The goal of SEM is to account for as much variation (covariation) between the variables as possible. Furthermore, in contrast to classic regression analysis, this methodology is able to consider measurement errors of variables (e. g., items). SEMs use the covariance matrix of all variables as input for an iterative fitting process. The estimation is done simultaneously for all relations that are represented in the model. To this end a statistical program (e. g., Analysis of Moment Structures (AMOS)) refines the estimates of the free parameters in the model in an iterative fitting process. The process stops when the statistical software is not able to minimize the residual matrix any further. The literature suggests a multitude of fit measures and indices to determine how good the covariance matrix of the hypothesized model fits the sample covariance matrix. The more variation and covariation the model is able to explain, the better. This thesis relies on Maximum Likelihood Estimation (MLE) to fit the model.

An important feature of SEMs is the ability to account for interdependencies between explanatory variables. Often one has to account for 'mediating effects'. Mediating variables are influenced by other explanatory variables in the model and have a direct effect on the explained variable. But there might not be a direct effect of the variable influencing the mediator on the explained variable. For instance, one could assume that the number of problems with the Internet connection directly influences the consumers' satisfaction with the status quo. However, there might not be a direct effect of the problems with the connection quality on the decision for IU tiering. Therefore, this variable would have only an indirect effect on the decision for IU tiering through its mediator 'satisfaction with the status quo'.

The aim of the following sections is to explain in detail the process of developing and validating the underlying model and hypotheses, as well as the design of the survey to collect the necessary data. Finally, the process of data evaluation and hypothesis testing is explained. Figure 4.2 shows the order of steps that have been taken to systematically develop and conduct this study.

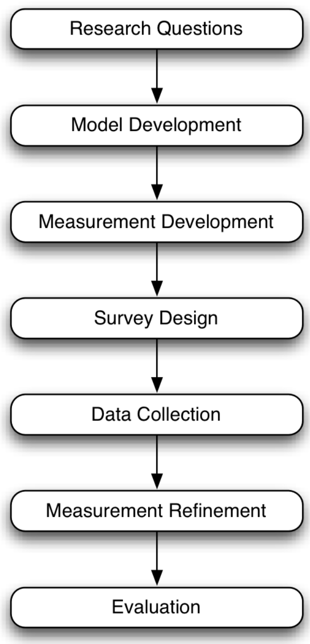

Figure 4.2.: Study procedure

4.3. Model

First, a research model has to be specified, that allows to test specific hypothesis in relation to the research questions. Based on this model, the survey can be designed with appropriate measures, questions and choice scenarios to gather the necessary data.

The model presented in this thesis relies on variables which have been deducted from existing literature and which are assumed to have an influence on the decision of the respondents in the hypothetical choice scenarios about an IU tiering scenario. In SEM one distinguishes between the measurement model and the structural model. The measurement model describes the mapping of items to the unobservable variables, whereas the structural model describes the assumed relations (i. e., hypotheses) between all variables in the model.

The measurement model employed in this thesis is assumed to be reflective. Reflective measurement models, in contrast to formative measurement models, are based on the assumption that the set of observable indicators (items) is influenced by the underlying unobservable latent variable and not the other way around (e. g., Zinnbauer and Eberl, 2004). These kind of measurement models do not assume that the final set of items, constituting the construct, covers necessarily all aspects of the latent variable. In contrast to that, a formative construct is constituted by a set of items that is assumed to incorporate all relevant aspects of the unobservable variable. In this case the items would have a direct influence on the latent variable, but are not necessarily correlated with each other (e. g., Ronald and Raschke, 2007). The items of a reflective construct should ideally be highly correlated, because they are to some extend interchangeable indicators of the same unobservable variable (e. g., Zinnbauer and Eberl, 2004).

The structural model is visualized by a path diagram. Each arrow in the diagram symbolizes a hypothesis about the relationship between the respective variables. Figure 4.3 shows a simplified representation of the structural model that is employed to study the effect of the aforementioned dimensions on the decision about hypothetical IU tiering scenarios. Note that the decisions of participants w. r. t. relative and dedicated quality provision are estimated simultaneously. The figure is simplified w. r. t. the granularity of the paths and dependencies between variables. To this end the variables are grouped in several boxes. Note that every arrow from one box to another box implies a single arrow from each variable in the originating box to each variable in the target box. All grey shaded variables are assumed to have an influence on the target variables in the model. According to the SEM notation convention, ellipses denote unobservable variables and rectangles denote observed variables. Double headed arrows symbolize that these variables are allowed to covary with each other (i. e., they have a relation outside the scope of the model). Single headed arrows, on the other hand represent a unidirectional relation. Therefore, each single headed arrow in the structural model is a hypothesis about the relationship of the two connected variables. It is obvious that not all relations necessary to account for all possible interactions of variables are in the core interest of this study. The extensive model specification can be found in Appendix A.2.

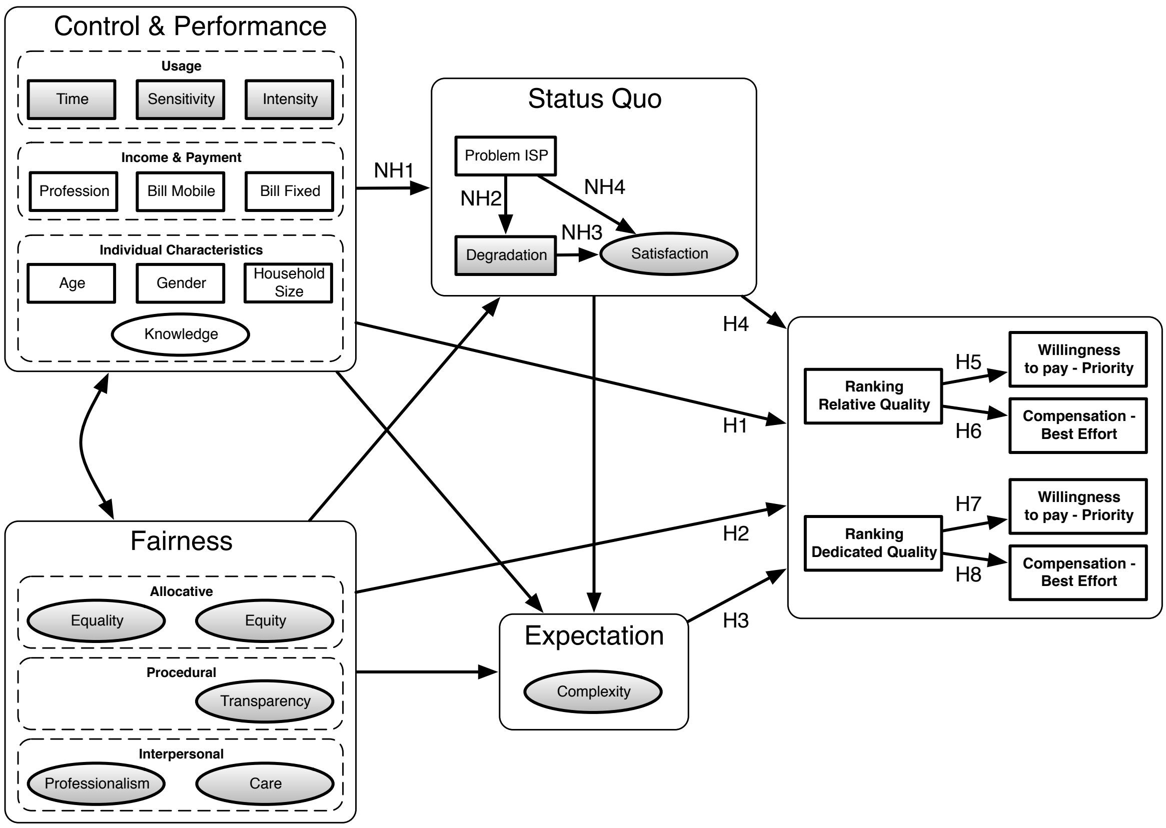

Figure 4.3.: SEM structure

The aim of this model is to explain variables that are in direct relation to the research questions motivated in this Chapter. Respondents are asked to make three choices. The first between NN and IU tiering with relative quality provision mechanism, the second between NN and

IU tiering with dedicated quality provision, and the third between IU tiering with relative and dedicated quality provision. These three choices allow to calculate a ranking of each quality provision mechanism for each participant. Consider as an example the variable *Ranking Relative Quality*. The variable takes values from one to three with higher values indicating that the respondent has a higher preference for IU tiering with a relative quality provision mechanism. Table 4.1 shows the ranking of the regimes and the assigned values.

Furthermore, respondents are asked to indicate their WTP for priority under each quality provision mechanism and to indicate the monetary compensation that would be necessary to make them accept Internet access service with BE quality, where they would suffer from the adverse effect of prioritization.

Table 4.1.: Network regime ranking variable

Preference	**Ranking Relative**	**Ranking Dedicated**
$Relative \succ Dedicated \succ NN$	3	2
$Dedicated \succ Relative \succ NN$	2	3
$Relative \succ NN \succ Dedicated$	3	1
$Dedicated \succ NN \succ Relative$	1	3
$NN \succ Relative \succ Dedicated$	2	1
$NN \succ Dedicated \succ Relative$	1	2

4.3.1. Control and Performance

The box 'Control and Performance' contains all variables that are assumed to influence all other variables in the model. They are necessary to account for heterogeneity in the sample. IUs differ w. r. t. their consumption pattern, expenditures, demographics and technological knowledge.

The usage pattern of an IU is captured by three variables. IUs are asked to indicate the average *Time* they spend on in the Internet per month. To capture detailed usage behavior, the survey presents six representative Internet service categories (derived from Sandvine, 2010), and respondents are asked to indicate their usage intensity on a five point Likert scale (0=services not used, 4=services frequently used).

1. Internet telephony (e. g. , VoIP) [sensitive]
2. Online gaming [sensitive]
3. Realtime entertainment [sensitive]
4. E-mail [insensitive]
5. Social networks [insensitive]
6. File sharing (e. g. , P2P) [insensitive]

The sum of all Likert points an IU indicates for these six categories is called usage *Intensity*. *Sensitivity* is captured by the share of an IU's congestion sensitive usage intensity. The variable relies again on the aforementioned six representative service types and is calculated by:

$$Sensitivity = \frac{\sum_{i=1}^{3} Usage_i[sensitive]}{\sum_{i=1}^{6} Usage_i}$$

The reason for choosing this observable measure is quite intuitive. Unified Theory of Acceptance and Use of Technology (UTAUT) suggest 'usefulness' as a measure for existing IT systems. In contrast to that IU tiering is a purely hypothetical scenario, and therefore it is difficult for IUs to evaluate the usefulness ex-ante, especially if they are not provided with specific performance gains they can expect from the system. However, users with a higher usage intensity of congestion sensitive services are very likely to expect higher gains from IU tiering. Based on this reasoning, the decision was made to include rather the objective demand for quality as an observable variable instead of an indirect measure. Thus *Sensitivity* is interpreted as observable economic representations of this aspect. This need is assumed to positively influence the ranking of IU tiering, as well as the willingness to pay for priority.

To control for the expenditures of IUs, participants are asked to indicated their average monthly bills for fixed and mobile Internet access (*Bill Mobile* and *Bill Fixed*). To account for possible income effects a proxy variable called *Profession* is used. Participants are asked to select from a list the job category they belong to. The variable *Profession* assigns a value from one to six to each category with increasing average income.

Individual characteristics of Internet access decision makers are controlled by *Age*, *Gender* and *Household Size*. *Knowledge* of computer systems and networks possessed by an IU is incorporated as a latent variable. Note that the participants' age is surveyed with an accuracy of 10 years. An increase of the age variable by one indicates that the respondent belongs to the next higher age group that is presented in the survey.

Hypothesis 1. *Users with a higher congestion Sensitivity evaluate IU tiering more favorable.*

> **Hypothesis 1.1.** *Users with a higher congestion Sensitivity have a higher preference for IU tiering with relative quality provision.*
>
> **Hypothesis 1.2.** *Users with a higher congestion Sensitivity have a higher preference for IU tiering with dedicated quality provision.*
>
> **Hypothesis 1.3.** *Users with a higher congestion Sensitivity have a higher WTP for priority under IU tiering with relative quality provision.*
>
> **Hypothesis 1.4.** *Users with a higher congestion Sensitivity have a higher WTP for priority under IU tiering with dedicated quality provision.*
>
> **Hypothesis 1.5.** *Users with a higher Internet consumption (Time and Intensity) have a higher WTP for priority under IU tiering with relative quality provision.*
>
> **Hypothesis 1.6.** *Users with a higher Internet consumption (Time and Intensity) have a higher WTP for priority under IU tiering with dedicated quality provision.*

4.3.2. Fairness

The box labeled 'fairness' contains constructs to account for the importance of different aspects in the context of fairness. Each construct measures how important participants in the survey perceive the different aspects of fairness in the context of Internet access services. Depending on these factors one can draw conclusions about how consumers decisions and expectations are driven by fairness. As explained before, one distinguishes between distributive, procedural and interactional justice. In the context of the NN debate, the most important aspects w. r. t. distributive justice are *Equality* and *Equity*. In other words, is it perceived as fair that every customer gets the same quality (equality), or that customers with a higher WTP get a better transmission quality (equity). These competing effects are also at the heart of the NN debate, because the transition from an BE network regime to IU tiering can be thought of as a transition from an equality driven network regime to an equity driven network regime. However, it is not clear ex-ante whether IUs perceive IU tiering as a violation of *Equality*, or as support of *Equity*.

Procedural justice is captured by the aspect of *Transparency*. This construct captures how important control and information about the Internet access service is to the consumer. Equity theory suggests that the perception of fairness depends on the procedure that determines the outcome. In the context of NN, regulators seem to be convinced that transparency is a key element to ensure that consumers can evaluate the quality provision and make informed decisions if they feel they do not receive the expected level of service. This aspect overlaps with the requirement of feedback mechanisms that has been identified by QoS studies (e. g. , Bouch and Sasse, 1999). Therefore, this construct also captures the importance of control over the network quality and functionality.

In the context of Internet access the aspect of interactional justice is captured by the importance of *Trust* in the ISP. This concept can be further subdivided into *Professionalism* and *Care*, both of which are key elements of trust. Like *Equality* and *Equity* these two constructs capture two partially competing effects. While *Professionalism* captures the importance of the ability and integrity of the ISP as a business partner, *Care* captures the importance of the impression of 'being in good hands' (i. e. , the benevolence) (cf. Xia et al., 2004, p.5). The interpretation of these two aspects can be ambiguous. If an event is assumed to be perceived as positive, the two aspects of *Trust* are assumed to work in the same direction. If a consumer has a positive experience with the ISP he may think that the ISP is professional and cares about him. A different result is expected with regard to negative events. For instance, if a consumer is legitimately (e. g. , for contractual reasons) not granted an upgrade into a better service class, he may still believe that the ISP is professional, but may doubt that he cares about him as a customer. In this case *Professionalism* works as "[...]a buffer that helps decrease negative attributions when price discrepancies occur"(Xia et al., 2004, p.9). In contrast to that, "[..]loyal customers focus more on whether the sellers care about them. When customers perceive an unfair price, they are likely to perceive it as exploitation and are more likely to punish the seller"(Xia et al., 2004, p.9).

The effect of fairness on the perception of IU tiering is the main reason for conducting this research project. The study of two general concepts of providing quality has been made because it is assumed that IU tiering is not evaluated based on the simple fact that quality is differentiated, but also on how this outcome is achieved. Relative priority provision may rely more on trust in the ISP, because it guarantees preferential treatment to paying customers, without making any explicit guarantees or capacity reservations in the network. By contrast, dedicated quality mechanisms may be perceived as more transparent, because they assign a certain amount of capacity solely to priority customers. In addition to that, prioritization based on relative priority ($M/M/1$) explicitly violates the equality principle, while splitting capacity in two service classes leads to two networks in which the same set of rules apply. Therefore the following hypotheses are made:

Hypothesis 2. *Fairness has an impact on the ranking of IU tiering, on the WTP for priority service and on the necessary monetary compensation to accept BE service.*

Hypothesis 2.1. *IUs who perceive Equality as an important concept in Internet access services have a lower preference for IU tiering with relative quality provision.*

Hypothesis 2.2. *IUs who perceive Equity as an important concept in Internet access services have a higher preference for IU tiering with dedicated quality provision.*

Hypothesis 2.3. *IUs who perceive Professionalism as important in their relation to the ISP have a higher preference for IU tiering with relative quality provision.*

Hypothesis 2.4. *IUs who perceive Care as important in their relation to the ISP have a higher preference for IU tiering with relative quality provision.*

Hypothesis 2.5. *IUs who perceive Transparency as important in Internet access provision have a higher preference for IU tiering with dedicated quality provision.*

Hypothesis 2.6. *IUs who perceive Equity as an important concept in Internet access services have a higher WTP for priority service under IU tiering with relative quality provision.*

Hypothesis 2.7. *IUs who perceive Equity as an important concept in Internet access services have a higher WTP for priority service under IU tiering with dedicated quality provision.*

Hypothesis 2.8. *IUs who perceive Professionalism as important in their relation to the ISP demand a lower monetary compensation for BE service under IU tiering with relative quality provision.*

Hypothesis 2.9. *IUs who perceive Professionalism as important in their relation to the ISP demand a lower monetary compensation for BE service under IU tiering with dedicated quality provision.*

Hypothesis 2.10. *IUs who perceive Care as important in their relation to the ISP demand a higher monetary compensation for BE service under IU tiering with relative quality provision.*

Hypothesis 2.11. *IUs who perceive Care as important in their relation to the ISP demand a higher monetary compensation for BE service under IU tiering with dedicated quality provision.*

4.3.3. Effort Expectancy

According to UTAUT, effort expectancy is an important dimension in the adoption process of new technologies. In contrast to the performance expectancy, effort expectancy cannot be captured by an observable proxy and is therefore incorporated by the perceived *Complexity* of IU tiering. Compared to the homogeneous service under a neutral BE network regime, the introduction of QoS and multiple classes of service could complicate the usage of the Internet from a customer's point of view. Furthermore, UTAUT shows that the expected *Complexity* of a new IT system is influenced by a user's general level of technical knowledge. To control for this factor, and to be in line with the underlying theory, *Knowledge* has been added to the model specification as exemplified in the control and performance section.

Effort expectancy is surveyed exclusively for the case of giving priority to some users. Therefore, effort expectancy is intended to capture the utility offset that comes along with the, in general, beneficial option to get preferential treatment. The construct does not capture any expectations about the BE lane under IU tiering to prevent any mixture with the aspect of performance expectancy, and because BE users do not have to make any additional effort to get in the BE lane, if they decide against priority service.

However, the perception of fairness and *Complexity* of a new network regime is most likely highly correlated with each other. The model can not explain the (unobservable) reasons why users perceive certain aspects of fairness as more important than others in the context of Internet access services, but these reasons are likely to implicitly influence the perception of *Complexity* as well.

Recall that the effort expectancy is expected to measure the utility offset of prioritization. Consequently no hypotheses about the impact on the necessary monetary compensation under BE are made.

Hypothesis 3. *IUs who perceive IU tiering as more complex evaluate IU tiering less favorable.*

> **Hypothesis 3.1.** *IUs who perceive IU tiering as more complex have a lower preference for IU tiering with relative quality provision.*
>
> **Hypothesis 3.2.** *IUs who perceive IU tiering as more complex have a lower preference for IU tiering with dedicated quality provision.*
>
> **Hypothesis 3.3.** *IUs who perceive IU tiering as more complex have a lower WTP for priority under IU tiering with relative quality provision.*
>
> **Hypothesis 3.4.** *IUs who perceive IU tiering as more complex have a lower WTP for priority under IU tiering with dedicated quality provision.*

4.3.4. Status Quo

The box 'status quo' incorporates the respondent's *Satisfaction* with his situation in the status quo, which is unobservable and therefore incorporated as a latent variable. The perception of the status quo is assumed to depend on possible past experience with problems of the Internet access service (*Problem ISP*) and the presumption to be artificially degraded by the ISP (*Degradation*). Again, *Problem ISP* relies on the six aforementioned service categories. Respondents have to indicate if they experienced at least once problems with the respective service class and, in addition, if they attribute this problems to the ISP. The relative proportion of the usage intensity that was at least once affected by problems attributed to the ISP is calculated by:

$$ProblemISP = \frac{\sum_{i=1}^{6} Usage_i[if\, problem\, attributed\, to\, ISP]}{\sum_{i=1}^{6} Usage_i}$$

Furthermore, respondents are asked to indicate for each individual service class if they assume to be artificially degraded by their ISP. Therefore *Degradation* is calculated analogously by:

$$Degradation = \frac{\sum_{i=1}^{6} Usage_i[if\, assumed\, to\, be\, degraded]}{\sum_{i=1}^{6} Usage_i}$$

In section 4.1 it has been exemplified that IUs compare alternative outcomes with their current outcome. Therefore, it is very likely that IUs infer from their past experience as well as current satisfaction about IU tiering scenarios. For instance, users who assume to be degraded by their ISP could fear the dirt road fallacy under IU tiering.

Hypothesis 4. *IUs who are more satisfied with the status quo evaluate IU tiering less favorable.*

> **Hypothesis 4.1.** *IUs who are more satisfied with the status quo have a lower preference for IU tiering.*
>
> **Hypothesis 4.2.** *IUs who presume more often to be degraded under the status quo have a lower preference for IU tiering.*
>
> **Hypothesis 4.3.** *IUs who presume more often to be degraded under the status quo demand a higher monetary compensation for BE service under IU tiering.*

In addition to these main hypotheses one can establish and test sub-hypotheses about the status quo itself. In general ISPs seem to not be very concerned about the effect of artificial quality degradation on their customer base. Presumably, they assume that many customers will attribute a negative experience with the service to other factors (e. g. , technical problems), or that they will not realize any restrictions at all.

Hypothesis 4.4. *IUs with a higher congestion Sensitivity are more likely to presume to be artificially degraded by the ISP.*

Hypothesis 4.5. *IUs who attribute more problems with the Internet access service to the ISP are more likely to presume to be artificially degraded by the ISP.*

Hypothesis 4.6. *IUs who presume more often to be artificially degraded by the ISP are less satisfied with the status quo.*

Hypothesis 4.7. *IUs who attribute more problems with the Internet access service to the ISP are less satisfied with the status quo.*

4.3.5. Target Variables

Under the respective network regime, the preference (ranking) of a procedure to provide QoS is assumed to have a positive impact on the WTP for prioritized Internet access and conversely, a negative effect on the necessary monetary compensation for BE service.

Hypothesis 5. *IUs who have a higher preference for IU tiering with relative quality provision have a higher WTP for priority under IU tiering with relative quality provision.*

Hypothesis 6. *IUs who have a higher preference for IU tiering with relative quality provision demand a lower monetary compensation for BE under IU tiering with relative quality provision.*

Hypothesis 7. *IUs who have a higher preference for IU tiering with dedicated quality provision have a higher WTP for priority under IU tiering with dedicated quality provision.*

Hypothesis 8. *IUs who have a higher preference for IU tiering with dedicated quality provision demand a lower monetary compensation for BE under IU tiering with dedicated quality provision.*

All hypotheses mentioned above are in the core of this study. Recalling Figure 4.3 reveals that the model contains also other relations between variables that are not explicitly mentioned above. In this respect the model is exploratory in nature. In general, SEMs face a trade-off between parsimony and explanatory power. Usually it is recommended that a model should only incorporate parameters and relations between parameters that help to explain the variation in the sample. However, the decision was made to refrain from ad-hoc assumptions about and omittances of relations between variables.

4.4. Measurement Development and Pretest

In preparation of the main survey, pretesting is an indispensable prerequisite to generate a suitable survey and suitable constructs for latent variables. First, items for each aforementioned unobservable variable were generated with an deductive approach according to aspects mentioned in the related literature. Alternatively existing constructs from the literature were adapted.

In total, the pretest consisted of 68 initial items that were assumed to match on 12 hypothetical unobservable variables. In the pretest stage of this study aspects of different fairness theories where considered (e. g. , scarcity and competition), as well as an extensive item set to measure the satisfaction with the status quo.
To make judgment about the impact of fairness on the choice sets in the main survey one hypothetical choice about IU tiering from the main survey was part of the pretest. This allowed to check which of the constructs assumed to measure fairness have an impact on the decision.

The pretest questionnaire, conducted with 112 students of economics at the Karlsruhe Institute of Technology (KIT), contained the following dimensions:

Satisfaction - ISP (6 items)

The actual satisfaction of an IU with his ISP. Derived from: (Gerpott et al., 2001; Janda et al., 2001; Wang et al., 2001).

Satisfaction - Quality (9 items)

The actual satisfaction of an IU with the perceived quality of his Internet connection. Derived from: (Gerpott et al., 2001; Janda et al., 2001; Kim and Stoel, 2004; Shin, 2007).

Fairness - Trust (5 items)

The importance an IU attaches to *Professionalsim* and *Care*. Derived from: (Chiou, 2004).

Fairness - Transparency (6 items)

The importance an IU attaches to procedural justice. Derived from: (Faulhaber, 2010).

Fairness - Equality (8 items)

The importance an IU attaches to the equality aspect of distributive justice. Derived from: (Bolton et al., 2003; Martín-Ruiz and Rondán-Cataluña, 2008; Wagstaff, 1994; Xia et al., 2004).

Fairness - Equity (6 items)

The importance an IU attaches to the equity aspect of distributive justice. Derived from: (Bolton et al., 2003; Martín-Ruiz and Rondán-Cataluña, 2008; Wagstaff, 1994; Xia et al., 2004).

Fairness - Need (5 items)

The importance an IU attaches to the need aspect of distributive justice. Derived from: (Bolton et al., 2003; Martín-Ruiz and Rondán-Cataluña, 2008; Wagstaff, 1994; Xia et al., 2004).

Fairness - Scarcity (6 items)

The fairness perception of price increases due to scarcity of a physical good. Derived from: (Frey and Pommerehne, 1993).

Fairness - Competition (5 items)

The competitiveness of the ISP market as perceived by the IU. Derived from: (Hahn et al., 2007).

QoS - Complexity (6 items)

The Complexity of IU tiering as perceived by the IU. Derived from: (Teo et al., 1999; Venkatesh et al., 2003).

Control - Knowledge (6 items)

The IU's Knowledge about computer systems and network technologies.

All constructs are assumed to be reflective and were evaluated by means of Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA), as well as checked for validity and reliability. This procedure follows the suggested steps in the development of measures for use in survey questionnaires by Hinkin (1998).

Based on the predictive power of the different fairness constructs on the IU tiering choice, the decision was made to rely solely on justice theory as described before. Hence the other aspects of fairness (i. e., Competition, Scarcity) were excluded from the main survey. Need is a variant of allocative justice, but it showed no impact on the IU tiering choice and did not correspond to a relevant aspect of the potential transition from a BE network to a IU tiering regime. The answers concerning satisfaction with the ISP in most parts mirrored the answers concerning satisfaction with the Internet connection. Moreover, many participants indicated that they were not able to make that fine granular judgment about different aspects of their connection quality. Consequently, satisfaction was combined in one construct and the item set was reduced to capture only the main aspects (satisfaction with the ISP and satisfaction with the delivered connection quality). All trust items show the desired level of factor loadings and the procedure associates the trust items to the sub-classes of trust mentioned before. TR1, TR2 and TR4, as well as T3 and T5 form unique factors. Textual inspection reveals the procedure separates the items that are related to how people are treated and appreciated by the ISP from those items that are more related to the business related aspects of the ISP. The former set of items (TR1,TR2,TR4) is from now on called *Professionalism*, while the latter set of items (TR3,TR5) will be called *Care*. The remaining constructs were refined through item reduction w. r. t. internal consistency, reliability and validity (construct reliablity, Average Variance Explained (AVE), Fornell Larcker criterion (Fornell and Larcker, 1981) and Cronbachs alpha (Cronbach, 1951)). The final set of items that is incorporated in the main survey is:

Control - Knowledge (4 items)

KN1 - "I know a lot about the technical functionality of the Internet."

KN2 - "In my circle of friends and my family I am considered an expert for computers."

KN3 - "I have in-depth knowledge about computer networks and data transmission technologies."

KN4 - "I often need help with computer problems"

Status quo - Satisfaction (2 items)

SA1 - "I am satisfied with my Internet service provider"

SA2 - "I am satisfied with my Internet connection quality"

Fairness - Professionalism (3 items)

TR1 - "It is important to me that my ISP is honest.

TR2 - "It is important to me that my ISP is reliable.

TR4 - "It is important to me that my ISP is professional.

Fairness - Care (2 Items)

TR3 - "It is important to me that my ISP understands me.

TR5 - "It is important to me that my ISP takes care of me.

Fairness - Transparency (5 items)

TP1 - "I feel only comfortable using the Internet, if I am acquainted with all details of my Internet tariff and the respective connection quality."

TP2 - "It is important to me to have easy access to all information related to the provided network quality."

TP3 - "It is important to me that all information about the provided network quality are unambiguous and understandable."

TP4 - "I feel uncomfortable using the Internet, if information about my Internet access tariff and the respective connection quality are not verifiable."

TP5 - "I feel uncomfortable, if contractual and technical restrictions of my Internet access service are not revealed to me."

Fairness - Equality (5 items)

EA1 - "Only if every Internet user gets the same network quality I perceive this as fair."

EA3 - "Only uniform transportation of data makes sure, that all Internet users have the same prerequisites."

EA4 - "If network capacity is split up equally between all Internet users, this is fair."

EA7 - "Nobody should be privileged in data transportation in the Internet."

EA8 - "Nobody should be disadvantaged in data transportation in the Internet."

Fairness - Equity (4 items)

EI2 - "It is fair that users who are willing to pay a premium get a prioritized Internet connection.

EI3 - "I think it is unfair, if other users get a prioritized Internet connection, because they are willing to pay a premium.

EI4 - "If a user wants to pay less than other users, it is fair that he gets a lower connection quality.

EI6 - "Prioritization of data for a premium fee is unfair, because other users experience a lower connection quality because of that.

QoS - Complexity (4 items)

CO1 - "I think using the Internet is easier, if all services and websites are treated equally."

CO3 - "The usage of the Internet becomes difficult, if I have to consider the option to prioritize data."

CO5 - "I think it is clearer, if it is only allowed to provide the same quality for all services and websites."

CO6 - "The option to prioritize data makes the usage of the Internet cumbersome for me."

Note that these measures are further analyzed with the final data set. Detailed statistical evaluation of the final constructs is provided in Section 4.7.

4.5. Survey Design

The final survey design has to incorporate questions that are able to obtain all data that is necessary to run the presented research model and to test the related hypotheses. To this end the survey consists of four parts, as shown in Figure 4.4. First, respondents are asked to indicate their overall monthly usage time, their monthly expenditures for telecommunications services and their satisfaction with the status quo. To capture detailed usage behavior, the survey presents the six aforementioned representative Internet service categories (derived from Sandvine, 2010), and respondents are asked to indicate their usage intensity on a five point Likert scale. Along these service categories, it is additionally surveyed if the respondents suspect to have been degraded by the ISP at least once in the respective service category. The share of usage intensity that is suspected to have been degraded is used in the model. The same holds true for connection problems in each service category, that are attributed to the ISP.

In the second part, respondents have to choose between different network regimes. In total, they have to make three network regime choices and indicate two times their willingness to pay for priority and BE access to the network.
Research Question 4 deals with the potential difference between relative and dedicated quality. Participants have to decide twice between NN and the respective IU tiering scenario (relative or dedicated quality). After each decision the participants are provided with a reference price for Internet access under NN. The survey is conducted with a representative German sample,

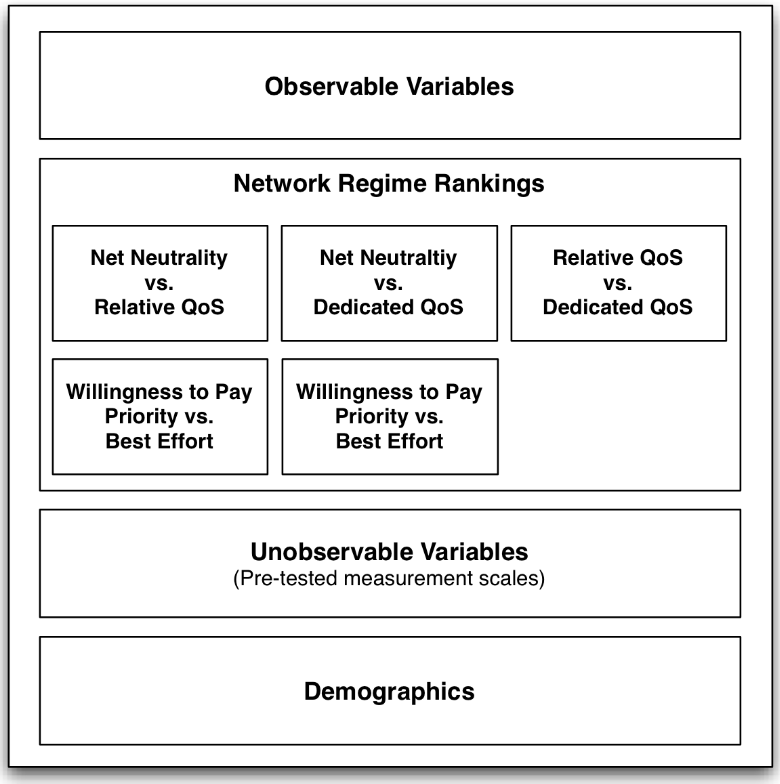

Figure 4.4.: Survey structure

where the lowest price for broadband Internet access is 20 Euros per month. Participants are asked to indicate their additional WTP for being in the priority service class under the respective IU tiering regime and their WTP for only being in the BE service. In the last question of this section participants have to choose between the relative and the dedicated mechanism of quality provision under IU tiering. All choice sets are explained in text form and visualized via an animated picture for better understanding. The pretest showed that the analogy of a toll (free) road helps participants best to gain understanding of the concept of IU tiering in relation to NN. Note that participants are advised to decide about Internet access and that they should consider the animation as an intuition helping them to simplify the complex underlying technical details. However, the survey never mentions the term NN, and all questions are phrased in a neutral language to ensure participants are not biased in favor of a certain network regime.

Figure 4.5 shows one frame of the animation that is used in the final survey. The upper part of the picture shows a representation of the NN regime, the lower part shows a representation of the IU tiering regime with relative quality provision.

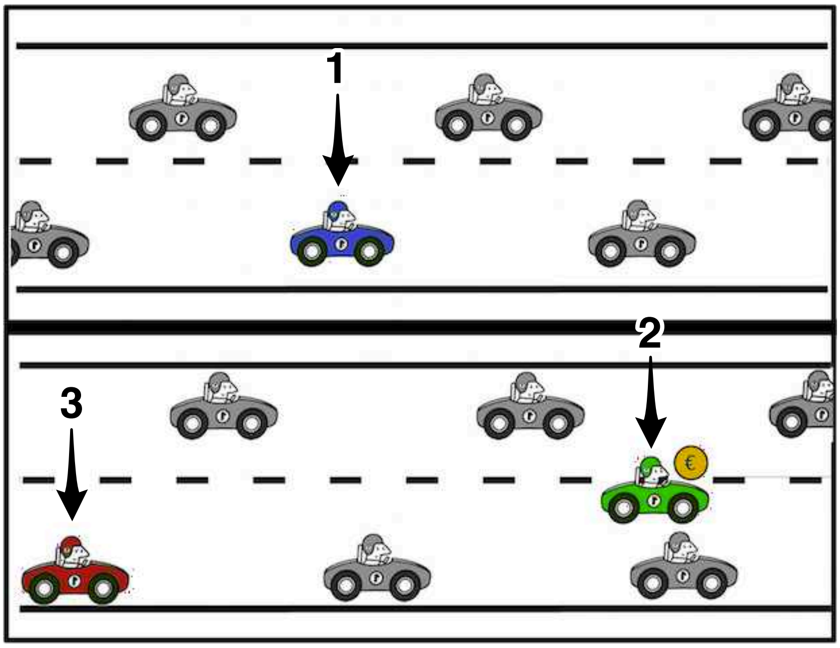

Figure 4.5.: NN vs. IU tiering relative quality

Relative quality is represented by an allusion to the $M/M/1$ queuing model (cf. Section 2.3.2). Each frame in the animation highlights three different cars to allow for easier comparison between the two regimes. In the web-based survey, only colors were used to highlight the respective car. Here they are additionally highlighted by numbers. All three highlighted cars start at the beginning of the road in each cycle of the animation. This allows the respondent to observe the improvement through prioritization and the degradation through the adverse effect of prioritization on the remaining users. The car labeled 'one' can be considered as the reference vehicle under a neutral network regime, whereas car 'two' has payed an additional fee (Euro symbol) and now has priority. The animation shows all other cars (including car 'three', a BE user) waiting until car 'two' (with priority) has passed by. Consequently car 'two' arrives first, followed by car 'one' and finally car 'three'.

This analogy is chosen in order to ensure that respondents with a non-technical background are able to compare the differences. This graphical representation is necessary, because it is absolutely essential that any participant understands the idea and implications of IU tiering, even if she/he has never heard of data prioritization before.

After participants have decided for one of the two network regimes two additional sliders appear and they have to indicate how much they are willing to pay for Internet access represented by car ‘two’ and with the second slider for Internet access represented by car ‘three’. The price for neutral Internet access is anchored at 20 Euros. The slider for priority only allows for higher prices up to 40 Euros, whereas the other slider only allows for lower prices down to zero Euros.

Figure 4.6 shows one frame of the animation for the second network regime decision. Participants have to decide between NN and IU tiering with dedicated quality provisioning. In contrast to the picture of the first decision, one lane of the street is dedicated only for paying users only. BE customers can only use the remaining and consequently more congested lane. This fact is represented by a slower pace of the cars in the BE lane.

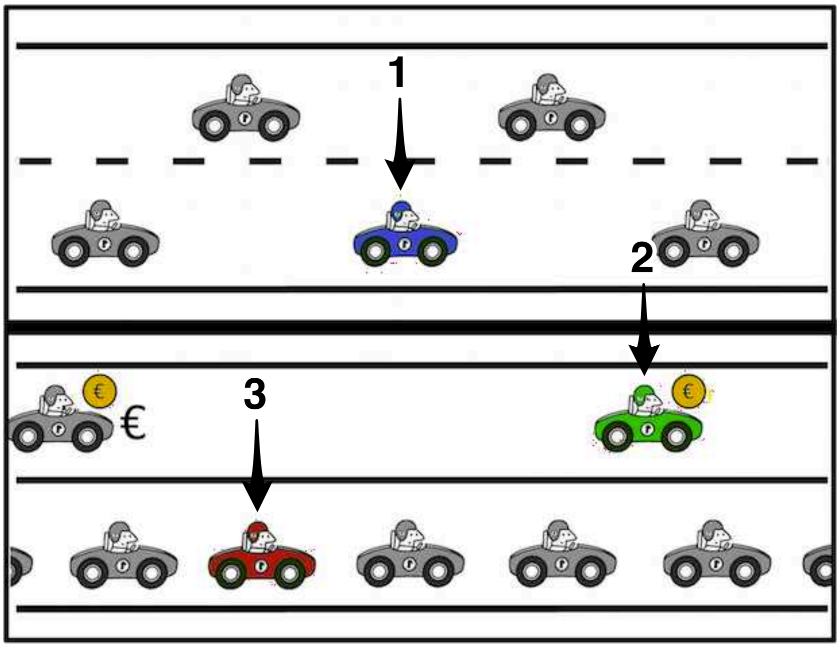

Figure 4.6.: NN vs. IU tiering dedicated quality

Again, after participants have decided for one of the two network regimes two additional sliders appear and they indicate their WTP for Internet access represented by car 'two' and, with the second slider, for Internet access represented by car 'three'. The slider behavior was identical to the first question.

Note that the two highlighted priority cars under IU tiering arrive at the same time at the end of the road under both quality provision mechanisms. This ensures that participants are not biased in favor of either network regime, because they expect any performance gains. The same holds true for the BE car under both quality provision mechanisms.

Figure 4.7 is the last network regime decision and simply compares the two kinds of quality provisioning under IU tiering.

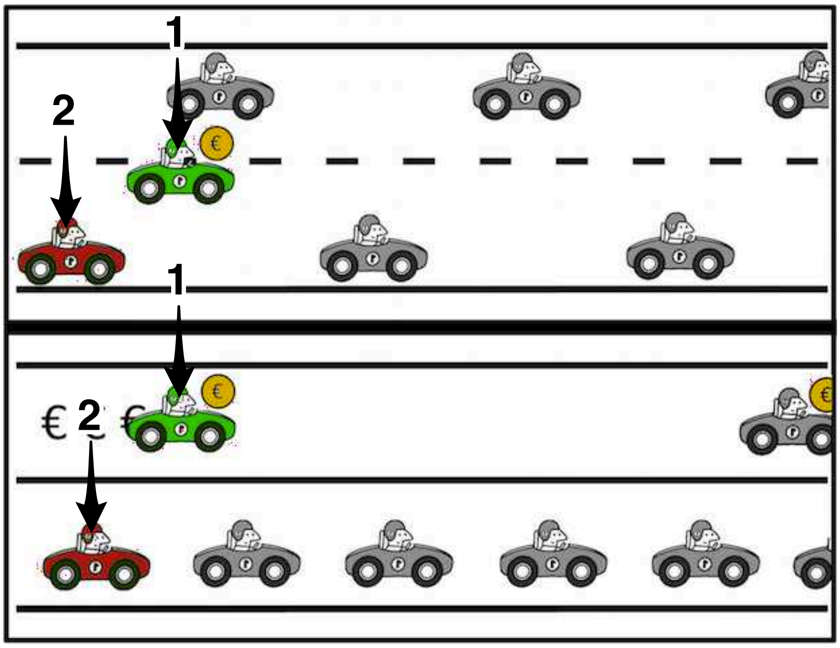

Figure 4.7.: IU tiering relative quality vs. IU tiering dedicated quality

In the third part of the survey, respondents are asked to indicate their level of agreement w. r. t. all items constituting the latent variables according to the pretest procedure. To reduce order effects, all items in the questionnaire are presented randomized to the participants.

In the last part participants are asked for usual demographics (cf. Section 4.3) and feedback, and finally whether they have been honest in filling out the survey.

The survey can be found in Appendix A.1.

4.6. Data Collection

The survey was conducted in June 2011 by a professional market research institute via an online panel among representative Internet access customers. The survey itself was implemented in the open source software solution 'Limesurvey'[3] and hosted by the Institute of Information Systems and Management (IISM). The questionnaire software allows the implementation of non-linear conditions in the online survey designs and provides access control mechanisms to ensure interoperability with professional panel providers. All participants were payed ex-post by the market research institute. They received a fixed participation fee and a variable compensation, depending on the median handling time of all participants in the survey. The market research institute did not reveal any payout information. However, the payout to each participant had to be below 2.9 euros, the fee (after tax) payed by the IISM to the market research institute for each participant. Participants were screened to ensure that they are the responsible persons in their respective household for choosing the Internet access provider and tariff. This screening also ensured that participants were familiar with the basic facts about consumer Internet access services and representative for the following hypothetical choice scenarios. If this requirement was not fulfilled, the questionnaire was aborted and a screen-out message was displayed to the participant by the panel provider. A soft-start of the main survey with 50 participants revealed that the median handling time was 8 minutes below the median time measured at pretests conducted at the IISM. Consequently, another screening stage was implemented to ensure participants were reading the questions carefully and not just clicking through the survey to make money (Questions 5-7 in Appendix A.1). All screening questions

[3] www.limesurvey.org

could be correctly answered by all participants that read the companion text on the same page. Due to the very simple nature of the screening questions it was made sure that screen-outs happened only due to a lack of reading the instructions and not a lack of understanding, which would have biased the sample in favor of versed participants. According to the panel provider, the logged total screen-out ratio was 37 percent of all invited participants. It is technically not possible to determine the share of participants who were screened out due to the lack of responsibility for their own Internet access service and the share of screen-outs based on improper reading of the instructions. However, the usual screen-out ratio without additional reading control is about 20 percent according to the panel provider.
In total 1035 IUs completed the survey. Ex-post 5 observations were excluded from the study due to self reported dishonesty of the participants, and another 50 due to intransitive order of network regime ranking decisions. Another 5 observations were sorted out due to severe non-normality of data (i. e. , respondents reported only 5 points on all Lickert scale items), leaving 975 valid observations for the evaluation process.

4.7. Measurement Refinement

Before one can test the hypotheses based on the collected data one has to verify and, if necessary, refine the measurement model by means of EFA and CFA. The preliminary specification of the constructs has to be tested for convergent validity, discriminant validity and reliability. After the refinement, the final constructs are evaluated based on their measures for construct reliability, AVE and the Fornell Larcker criterion (Fornell and Larcker, 1981). To estimate the model, the statistical software package Statistical Package for the Social Sciences (SPSS) (Version 19) and the companion tool AMOS was used.

EFA is a statistical data dimension reduction procedure to find a smaller set of underlying summary variables (factors). Even if one has already created groups of items with the intention to measure a single latent variable, this step in the evaluation process is necessary to verify that the items indeed constitute a unique factor. This is possible because EFA does not assume any relations between the items ex-ante. According to Field (2000), EFA uses the correlation matrix of all items and groups those correlating strongly but weakly with all other items. Ideally, items designed to measure a single latent variable load predominantly on the same factor and relatively low on all other factors. The procedure assigns items to additional factors as long as an additional factor explains a sufficient part of the variance in the sample. Costello and Osborne (2005) suggest 'Principal Axis Factoring' as the best method to retain factors if data is not necessarily normally distributed. In this specific study one has to further assume that the five dimensions of fairness (*Professionalism*, *Care*, *Transparency*, *Equity* and *Equality*) correlate with each other and, consequently, these factors are not orthogonal to each other. Therefore, one has to use an oblique rotation method (e. g. , 'Promax') in the EFA that allows the rotation procedure to extract factors in a non-orthogonal position to each other (Field, 2000). This procedure yields more accurate results. Kaiser Normalization is a standard transformation procedure in EFA with the effect, that the squared factor loadings in each row sum up to one. The squared factor loading of an item shows the proportion of the variance of the item that is explained by the factor. Table 4.2 shows the resulting pattern matrix. The pattern matrix presents the 'pattern loadings', which can be interpreted as the regression coefficients of each item on each of the extracted factors (DeCoster, 1998). Values lower than 0.2 are suppressed for clarification. First, note that all items load predominantly on the intended factors, which is an indicator for convergent validity. Item 6 from the *Equity* construct (EI6), as well as items CO1 and CO5 show relatively high cross-loading with the factor equality (EA). This is problematic and has to be further analyzed by CFA.

In contrast to EFA, CFA assumes a fixed relation of items to latent variables. Constructs that correlate with each other have to be tested for discriminant validity (i. e. , how selective they are w. r. t. the correlated variables). This test gives statistical indication whether the constructs at hand really measure two distinct unobservable variables. Several goodness of fit indices and modification indices for CFA are available to detect problematic items that distort the accurate measurement of the latent variables. However, it is also very important to keep in mind

Table 4.2.: EFA - Initial pattern matrix

	Factor							
	1	2	3	4	5	6	7	8
EA4	.927							
EA1	.886							
EA7	.827							
EA8	.809							
EA3	.794							
TP4		.806						
TP1		.768						
TP3		.754						
TP5		.723						
TP2		.684						
KN2			.863					
KN3			.809					
KN1			.804					
KN4			.571					
CO3				.885				
CO6				.848				
CO1	.213			.597				
CO5	.248			.487				
EI2					,843			
EI4					,726			
EI3					,643			
EI6	-.343				,497			
TR2						.903		
TR1						.709		
TR4						.705		
TR5							.853	
TR3							.715	
SA1								.803
SA2								.776

Extraction Method: Principal Axis Factoring.
Rotation Method: Promax with Kaiser Normalization.

the verbal representation of each item and to inspect the statements again before items are removed from the study (content validity). Only if one can find evidence for contextual misinterpretation of items that supports the statistical analysis one should refine the constructs by omitting them. Otherwise this iterative process can easily divert to simple model tweaking.

Inspection of the confirmatory results shows that the regression weights for all items are significant. However, the overall model fit is insufficient ($CMIN/DF = 4.5$ and RMSEA $= 0.09$). Inspection of the modification indices and cross checking with the item texts reveal several problems that have to be resolved.

First, the constructs *Equality* and *Equity* are not sufficiently selective to each other. Consequently the items causing this had to be detected and removed. EA7 and EA8 show a high association with the *Equity* construct which can be explained by the explicit mentioning of the words ‘privileged’ and ‘disadvantaged’. EI2 has a kurtosis of above ten, which violates the boundaries to legitimately assume a normal distribution (e. g. , Kline, 2005). Removing these three items delivers discriminant valid constructs.
Second, item TP5 shows a very small regression weight (<0.5). Textual inspection of the item leads to the conclusion that this is the only item not mentioning ‘network quality’, but instead ‘contractual and technical restriction’. The item is therefore removed.
Third, item KN4 also has a borderline regression weight in the CFA. Textual inspection leads to the conclusion that this item suffers presumably from the well known ‘social desirability’ bias (Fisher, 1993). The item is therefore removed.
CO1 and CO5 show a very high association with the equality construct. Textual inspection reveals that the verbal negation of these two items lead to the use of equality related phrases. This most likely caused the lack of discriminant validity. The two items are therefore removed.

For the final set of items an additional EFA is conducted. Note that after the refinement all items in Table 4.3 load predominantly on the intended factors (convergent validity) and show a sufficient level of factor loadings. Even though this result is desirable, one has to verify its

Table 4.3.: EFA - Final pattern matrix

	Factor							
	1	2	3	4	5	6	7	8
TP4	.807							
TP1	.767							
TP3	.757							
TP5	.724							
TP2	.681							
EA4		.941						
EA1		.854						
EA3		.692						
KN3			.833					
KN2			.825					
KN1			.824					
TR2				.925				
TR1				.708				
TR4				.700				
EI3					.946			
EI6					.684			
EI4					.509			
CO3						.869		
CO6						.832		
TR5							.809	
TR3							.773	
SA1								.792
SA2								.787

Extraction Method: Principal Axis Factoring.
Rotation Method: Promax with Kaiser Normalization.

validity to retain the obtained number of factors from the data. The following explanation of determining this number of factors is based on the article of Brown (2001). In factor analysis, eigenvalues are used to inspect for how much of the overall variance a factor accounts in a correlation matrix. For a given factor, the sum of the squared factor loadings is the eigenvalue. The scree test is used to find the 'kink' where the level of eigenvalues appears to level off. To the right of this point one can assume to find only 'scree'. Scree is the geological term referring to the accumulation of fragments at the base of mountains. It means that one should not extract too many factors, and that all factors to the right of the kink should thus be omitted. Note, that no other formal criterion exists to interpret the scree plot. The scree plot in Figure 4.8 suggests that the number of factors that have been extracted is valid.

Table 4.4 shows that the eight retained factors account in total for $76,54\%$ of the variance. Before the refinement procedure the factors were only able to account for 71.8% of the variance.

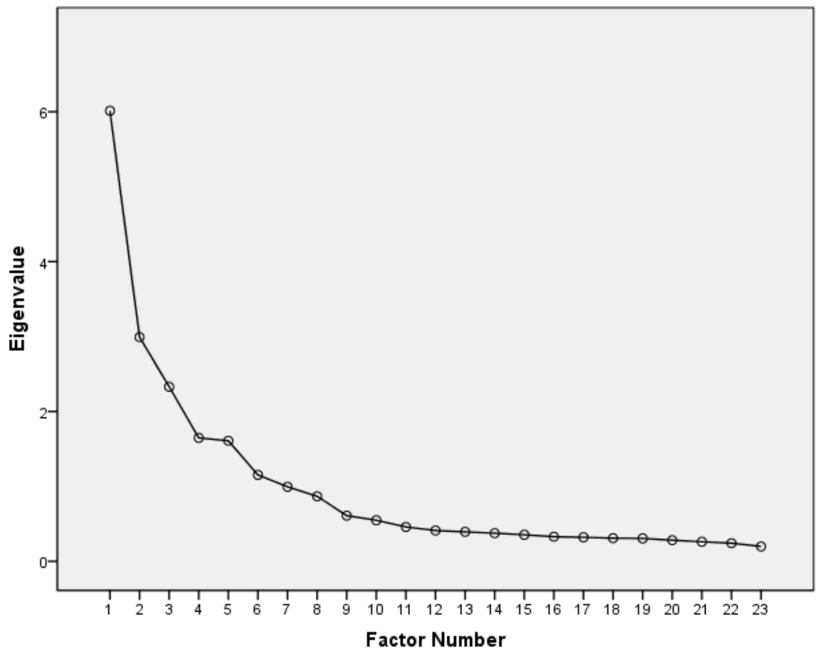

Figure 4.8.: Scree plot

Table 4.4.: EFA - Total variance explained

		Initial Eigenvalues		Rotation Sums of Squared Loadings[a]
Factor	Total	% of Variance	Cumulative %	Total
1	6.013	26.142	26.142	4.287
2	2.992	13.010	39.151	3.940
3	2.329	10.126	49.277	2.118
4	1.647	7.163	56.440	3.579
5	1.609	6.996	63.436	3.287
6	1.153	5.013	68.449	2.708
7	0.994	4.324	72.773	2.751
8	0.867	3.770	76.543	1.272

Extraction Method: Principal Axis Factoring.
a. When factors are correlated, sums of squared loadings cannot
be added to obtain a total variance.

To check the final constructs for discriminant validity and reliability, a second CFA is conducted. Assessment of normality reveals that it is valid to use Maximum Likelihood (ML) estimation because all items show a skewness below three and a kurtosis below ten (e. g., Kline, 2005).

All items load significantly on their latent variable. The overall measurement model shows a good model fit ($CMIN/DF = 2.1$, RMSEA = 0.034). Table 4.5 reports the standard measures for all latent constructs. Cronbachs alpha (Cronbach, 1951) is a measure for the internal consistency of a construct. If the correlations between the items of a construct are high, the measure delivers higher values. Cronbachs alpha attains values between zero and one. The literature suggests different thresholds for acceptable values. The most strict threshold (0.7) is postulated by Nunnally (1967, p.245), and all retained constructs fulfill this requirement.

The measure for factor reliability and the AVE (Fornell and Larcker, 1981) indicate how good the items of a construct are able to measure the latent variable. These measures are very sensitive to high measurement errors of the items. They are indicators for reliability and convergent validity of constructs (Zinnbauer and Eberl, 2004).

Discriminant validity can be checked by a simple χ^2 difference test. However, this thesis relies on the more sophisticated Fornell Larcker criterion (Fornell and Larcker, 1981). The criterion requires every factor to account for more variance than the squared highest correlation ($\max \rho^2$) with other factors. This ensures that only those constructs are included in the analysis that are able to explain enough variance on their own, even if they are allowed to correlate with each other.

Based on the results from Table 4.5, it is valid to proceed with hypotheses testing based on the proposed structural model using the final constructs as measures for the unobservable latent variables.

Table 4.5.: CFA - Construct reliability

Construct	**Items**	**Chronbachs α** ($\geq 0,7$)	**Reliability** ($\geq 0,7$)	**AVE** ($\geq 0,5$)	**Fornell Larcker** (AVE $> \max \rho^2$)
Equality	3	0,89	0.89	0.73	$0.73 > 0.50$
Equity	3	0.78	0.79	0.56	$0.56 > 0.50$
Care	2	0.79	0.79	0.66	$0.66 > 0.40$
Professionalism	3	0.84	0.80	0.58	$0.58 > 0.40$
Transparency	4	0.83	0.83	0.56	$0.56 > 0.24$
Complexity	2	0.84	0.84	0.72	$0.72 > 0.24$
Knowledge	3	0.86	0.86	0.68	$0.68 > 0.01$
Satisfaction	2	0,76	0.73	0.58	$0.58 > 0.01$

4.8. Evaluation

The model is estimated with the statistical software package AMOS. AMOS is especially designed for modeling and estimating SEMs. This section provides stepwise estimation output necessary to analyze all hypotheses.

4.8.1. Model Fit

Before one can start with the inspection of the results, the model fit has to be determined and evaluated. To this end AMOS reports a multitude of fit indices. The SEM literature is not always conclusive about the most suitable thresholds. To better understand the measures and their implication for this model, several fit indices are reported and discussed. First χ^2 tests the null hypothesis that the proposed (overidentified) model fits the data as good as a saturated (just-identified) model. A saturated model contains as many parameter relations as possible, only constrained by the degrees of freedom, whereas the the independence model constrains all relations between the parameters to zero. Table 4.6 reports the CMIN which represents the χ^2 statistics according to the number of estimated parameters in the model (NPAR). The CMIN alone is not suitable to evaluate the model fit, because it is very sensitive to sample size (p-value indicates that the null hypothesis of good model fit should be rejected). Therefore, CMIN/DF has been developed as a suitable alternative. Smaller values of the χ^2 statistics divided by the degrees of freedom indicate a better model fit. The most conservative threshold CMIN/DF$\leq$2 is suggested by Byrne (1989, p.55).

Table 4.6.: Fit - CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	395	743.238	424	0.000	1.753
Saturated model	819	0.000	0		
Independence model	78	15185.003	741	0.000	20.493

One of the most important fit measures in SEM studies ist the Root Mean Square Residual (RMSEA). The RMSEA is an indicator of how well the estimated model fits the population covariance matrix. It estimates the lack of fit to the saturated model. Consequently lower values are more desirable. The RMSEA favors more parsimonious models, because it is very sensitive to the number of estimated parameters. In addition, AMOS provides a confidence interval. The literature discusses many different cut-off values that indicate an acceptable fit. Browne and Cudeck (1993) suggest values below 0.08 and Hu and Bentler (1999) suggest values below 0.06. Table 4.7 reveals that the proposed model also fulfills this requirement and that the confidence interval also includes only desirable values.

Table 4.7.: Fit - RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	0.028	0.024	0.031	1.000
Independence model	0.141	0.140	0.143	0.000

So-called baseline comparisons are a group of measures that compare the proposed model with the independent model instead of the saturated model. Consider for instance the Normed Fit Index (NFI). The NFI is calculated as the difference between the χ^2s of the indepentent and proposed model divided by the χ^2 for the independence model. Table 4.8 reports also the Comparative Fit Index (CFI), Incremental fit Index (IFI) and Tucker Lewis Index (TLI). However, these measures are relatively similar to each other and all of them demand values above 0.90 to indicate acceptable fit and values above 0.95 to indicate good fit (Bentler and Bonett, 1989; Baumgartner and Homburg, 1996).

Table 4.8.: Fit - Baseline Comparisons

Model	NFI	IFI	TLI	CFI
Default model	0.951	0.978	0.961	0.978
Saturated model	1.000	1.000		1.000
Independence model	0.000	0.000	0.000	0.000

Appendix A.3 reports also the parsimony adjusted measures and the Akaike information criterion. Parsimony adjusted measures are variants of the NFI and CFI, that additionally penalize for model complexity. However, no general accepted cut-off values for these measures exist, which makes them difficult to interpret (Zinnbauer and Eberl, 2004). The Akaike information criterion is often used to compare competing models and to select the best fitting parsimonious model.

Due to its exploratory nature, the model might not be the most parsimonious one, but ex-ante it would be unjustified to remove paths from the model and make ad-hoc assumptions. In this respect the model is very conservative and controls for many variables that could influence the ranking of network regimes and the willingness to pay. Nevertheless, pretesting as well as solid model and survey design are rewarded by overall good fit values. One can therefore continue to inspect the estimations of the model and interpret the results. The following subsections focus predominantly on the presented Hypotheses and the interpretation of direct effects. However, the estimates of the remaining dependent variables, as well as indirect effects and covariances are reported in Appendix A.4.

4.8.2. Network Regime Preference

Table 4.9 shows side-by-side the effects of all variables on the network regime ranking under relative and dedicated quality provision mechanisms. IUs with a higher congestion *Sensitivity* are not more likely to prefer IU tiering or have a higher WTP for priority under either quality provision mechanism. This result can be interpreted in two different ways. Participants in the survey may not be fully aware of their consumption sensitivity, and therefore do not know that they would have additional gains from a better quality provision in Internet access services. One could also argue that they might be aware of their congestion sensitivity, but have a low valuation for additional quality, and therefore make decisions based on other criteria. Based on these results Hypotheses 1.1-1.2 have to be rejected. However, the effect on the dedicated quality provision mechanism is at least significant on the 7.5 percent level. Therefore, one could argue that dedicated quality provision is at least more likely to find acceptance among users that are aware of their congestion sensitivity, even if the potential effect is very small.

Fairness has indeed an impact on the network regime ranking of participants. In this respect participants perceive the two quality provision mechanisms fundamentally different from each other. Participants who perceive *Equity* as a fair concept in Internet access services are more likely to prefer the dedicated quality provision mechanism. In contrast to that, *Equity* does not significantly affect the preference for relative quality provision. Furthermore *Equality* has only a significant effect on the preference for the relative quality provision mechanism. However, the effect is only significant on the 7.5 percent level, but it is comparably high. This supports the non significant effect of *Equity* and indicates that relative quality provision mechanisms might be perceived as a violation of the equality principle. Therefore this mechanism might be more likely to be rejected by customers. Based on these findings Hypothesis 2.1 has to be rejected and Hypothesis 2.2 can be accepted. *Professionalism* has no measurable effect on the ranking of IU tiering with relative quality provision. Conversely, *Care* has an highly significant positive effect on the ranking of IU tiering with relative quality provision. The overall hypothesis was that relative quality provision is more appealing to IUs who perceive trust as an important aspect in the relation to the ISP. The fact that only *Care* has an effect may lead to the conclusion that IUs feel appreciated by the ISP through the possible assignment of priority. IUs that perceive it as important to feel treated special and with care are presumably more likely to rank a network regime higher that allows for such preferential and special treatment. Explicit prioritization taps this preference in contrast to simply splitting up capacity. Dedicated quality provision mechanisms, however, tap the preference for *Transparency* instead. In contrast to relative quality provision this mechanism relies less on trust and therefore IUs with a higher preference for procedural justice rank it higher. Based on these results Hypothesis 2.3 has to be rejected and Hypotheses 2.4 and 2.5 can be accepted.

A similar divergence can be found w. r. t. the *Complexity* of IU tiering. Perceived *Complexity* has only a significant negative effect on IU tiering under relative quality provision. This leads to the conclusion that participants in the survey perceive relative quality provision mechanisms as more complex and cumbersome, and therefore have a lower preference for this quality provision variant. This findings confirm Hypothesis 3.1 and leads to the rejection of Hypothesis 3.2.

The *Satisfaction* with the status quo has no significant effect on the ranking of both variants of IU tiering. One could interpret this finding as follows: The participants in this survey did not consider IU tiering as less useful even if they were already satisfied with the quality of the Internet access service and the ISP under the status quo. Therefore, one could conclude

Table 4.9.: Estimates - Ranking quality provision mechanism

	Relative QoS			Dedicated QoS		
	Estimate	St. Estimate	P	Estimate	St. Estimate	P
Complexity	**-0.104**	**-0.150**	***	0.035	0.043	0.323
Transparency	-0.053	-0.056	0.241	**0.148**	**0.133**	**
Care	**0.134**	**0.147**	**	-0.101	-0.096	#
Professionalism	-0.050	-0.034	0.520	0.086	0.050	0.358
Equity	0.064	0.069	0.225	**0.239**	**0.222**	***
Equality	-0.075	-0.102	#	-0.012	-0.014	0.797
Knowledge	0.003	0.004	0.912	0.019	0.020	0.612
Age	**0.006**	**0.077**	*	0.002	0.018	0.641
Gender	0.052	0.038	0.268	**-0.16**	**-0.100**	**
Sensitivity	-0.107	-0.029	0.410	0.296	0.068	#
Satisfaction	-0.007	-0.011	0.749	0.023	0.030	0.390
Degradation	0.064	0.032	0.334	0.027	0.011	0.735
Problem	-0.014	-0.006	0.852	0.063	0.023	0.481
Intensity	**0.019**	**0.095**	*	0.009	0.040	0.314
Bill Fixed	0.001	0.015	0.640	0.001	0.024	0.474
Bill Mobile	0.001	0.022	0.497	0.003	0.064	#
Time	-0.001	-0.003	0.918	-0.006	-0.025	0.463
Household	0.025	0.043	0.172	0.003	0.004	0.897
Profession	0.013	0.027	0.442	-0.002	-0.003	0.927

# p<0.075, * p<0.05, ** p< 0.01, *** p < 0.001

that consumers are at least open to additional quality gains that could be achieved by QoS mechanisms, even if they are satisfied with the status quo. Therefore Hypothesis 4.1 has to be rejected. Furthermore, it was assumed that users would project the negative experience they may had with artificial quality degradation in the status quo to an IU tiering scenario. They could assume that the ISP would try to artificially degrade the BE lane to lurk them into the costly priority lane. However, the results show that participants who assume more often to be degraded by their ISP are not significantly more likely to reject IU tiering. Therefore Hypothesis 4.2 has to be rejected.

Besides the explicitly mentioned hypotheses some control variables also show a significant effect on the preference for IU tiering. Inspection of demographics reveals that older participants are more likely to favor the relative quality provision mechanism, while males are less likely to favor dedicated quality provision. One can also observe that higher usage intensity has a positive effect on the ranking of the relative quality provision mechanism. Standardized estimates reveal that the main drivers for the ranking of the relative quality provision mechanism are *Complexity* and *Care*, while the main drivers for the ranking of the dedicated quality provision mechanism are *Transparency* and *Equity*.

4.8.3. Willingness To Pay for Priority

Table 4.10 shows the estimates for the variables that are assumed to have an influence on the willingness to pay for priority access under both quality provision mechanisms. First, note that consumption is a strong predictor of the WTP for priority under IU tiering under both quality provision mechanisms. This finding holds true for all variables that can be attributed to a high level of Internet usage, i. e., usage *Time*, usage *Intensity*, *Household* size and the expenditures for fixed and mobile access (*Bill Fixed*, *Bill Mobile*). Based on these findings Hypotheses 1.5 and 1.6 can be accepted.
Sensitivity does not have any effect on the WTP for priority under both quality provision mechanisms. The possible explanations for the non existing effect are identical to the ones for the non existing effect on the preference for IU tiering. Therefore Hypotheses 1.4 and 1.5 have to be rejected.

Fairness also has an impact on the WTP for priority. *Equity* has a highly significant effect under both QoS mechanisms which leads to Hypotheses 2.6 and 2.7 being confirmed. This finding may sound counterintuitive in the light of the diverging effects of fairness in the ranking. However, in fact payment decisions almost deterministically trigger the concept of equity fairness in humans. If people are confronted with a certain network regime, the WTP for additional benefits obviously does not depend on fairness concepts that may have influenced the preference for the regime itself. However, the ranking still has a significant effect on the WTP. The higher the ranking of a certain quality provision mechanism, the higher is the WTP for

Table 4.10.: Estimates - WTP Priority

	Relative QoS			Dedicated QoS		
	Estimate	St. Estimate	P	Estimate	St. Estimate	P
Ranking	**0.533**	**0.079**	***	**0.465**	**0.071**	***
Complexity	-0.205	-0.043	0.291	-0.346	-0.065	0.11
Transparency	0.416	0.064	0.167	0.578	0.079	0.086
Care	-0.173	-0.028	0.582	-0.214	-0.031	0.542
Professionalism	0.738	0.073	0.153	0.530	0.047	0.357
Equity	**1.374**	**0.220**	***	**1.882**	**0.266**	***
Equality	0.272	0.055	0.305	0.328	0.058	0.268
Knowledge	-0.216	-0.040	0.288	-0.236	-0.039	0.298
Age	-0.018	-0.033	0.365	**-0.056**	**-0.089**	*
Gender	**-0.779**	**-0.083**	*	-0.483	-0.046	0.165
Sensitivity	0.756	0.030	0.378	0.671	0.023	0.483
Satisfaction	0.047	0.010	0.752	0.207	0.040	0.217
Degradation	0.115	0.008	0.792	0.147	0.009	0.763
Problem	0.669	0.042	0.178	0.520	0.029	0.348
Intensity	**0.231**	**0.173**	***	**0.257**	**0.170**	***
Bill Fixed	**0.018**	**0.068**	*	**0.026**	**0.090**	**
Bill Mobile	**0.019**	**0.081**	*	**0.026**	**0.100**	**
Time	**0.140**	**0.109**	***	**0.107**	**0.074**	*
Household	**0.313**	**0.080**	**	0.246	0.056	#
Profession	-0.054	-0.016	0.633	-0.091	-0.024	0.471

# p<0.075, * p<0.05, ** p< 0.01, *** p < 0.001

priority under the respective quality provision mechanism. Standardized estimates reveal that the effect is relatively low compared to the direct effect of *Equity* and the aggregated effect of heavy consumption. Nevertheless, Hypotheses 5 and 7 are thereby confirmed. Some of the aforementioned fairness concepts that influence the ranking consequently still have an implicit effect on the WTP. These indirect effects are reported in Table A.7 in the appendix.

Effort expectancy shows no significant impact on the WTP and thus Hypotheses 3.3 and 3.4 have to be rejected.

Besides the explicitly hypothesized effects control for demographics reveal that older participants are willing to pay less for additional quality under relative quality provision and that males are less likely to pay more under dedicated quality provision.

In summary, one can conclude that heavy users are willing to pay more for additional quality of service, compared to the status quo. Fairness seems to play an important role, too. If users perceive *Equity* as a suitable concept in Internet access services, they have a significantly higher WTP.

4.8.4. Monetary Compensation for Best Effort

Table 4.11 shows the estimates for the variables that are assumed to influence the monetary compensation that would be necessary for consumers to accept BE Internet access under the respective quality provision mechanism.

It was assumed that users would project the negative experience they may had with artificial quality degradation to the BE service under IU tiering. IUs could assume that the ISP would try to artificially degrade the BE lane to lurk them into the costly priority lane. However, the results do not support Hypothesis 4.3 and therefore it has to be rejected.

The effect of fairness on the necessary monetary compensation is supported by the results. Participants who perceive *Care* as an important factor in their relation with the ISP are more likely to demand a higher monetary compensation for BE service. In contrast to that, *Professionalism* has a negative effect, which is only significant under relative quality provision, but at least significant on the 7.5 percent level under dedicated quality provision. However, the general intuition exemplified before seems to be confirmed by the data. In contrast to priority, BE triggers the feeling that the ISP might not care about BE customers, but this effect is partially counterweighted by the believe in its *Professionalism*. Even if consumers have to be compensated, because they are afraid that the ISP might not care about the BE class (or even feel betrayed), *Professionalism* makes it more likely that they demand less compensation under BE. The intuition behind this result can be compared to first and second class service in an airplane. Customers appreciate the better service in the first class, because they like to be taken care of personally and they can rely on the professional handling with a good airline anyway. In the second class, customers, who like to be taken care of might be afraid that the airline is not interested in their individual well being, and consequently they have also a lower

WTP for the economy class. But at the same time they know that they can rely on a certain quality of service, because the airline also handles the economy class with professionalism. Based on these results, Hypotheses 2.8, 2.10 and 2.11 are accepted and Hypothesis 2.9 has to be rejected.

The network regime ranking was expected to show a negative effect on the monetary compensation. This can only be confirmed for dedicated quality provision. Hypothesis 6 therefore has to be rejected and Hypothesis 8 can be accepted, but the effect size is comparably small. In addition to that, Table A.7 reveals that there are no significant indirect effects on the necessary monetary compensation under either quality provision mechanism. Therefore, in contrast to the WTP, the preference for IU tiering plays no decisive role w. r. t. the necessary monetary compensation under either quality provision mechanism. Besides the explicitly formulated hypotheses, one can observe four additional effects. First, *Knowledge* has a negative effect on the monetary compensation. IUs with a higher knowledge about computers and networks are more likely to demand a lower monetary compensation. Versed participants most likely assume the adverse effect of priority on other IUs to be comparably low, and consequently they are likely to accept BE for a higher price. Other IUs are most likely over-estimating the adverse effect and assume a higher disutility from BE service.

In contrast to that, higher usage *Time* makes it more likely to demand a higher compensation. Obviously users, that spend a lot of time on the Internet have to be compensated more under BE, because they expect a higher disutility due to the prevailing adverse effect of prioritization of other users.

Inspection of the demographic control variables reveals a highly significant and comparably big effect of age. Obviously older participants have to be compensated significantly more to accept BE service. This finding could indicate a lower price sensitivity of younger customers. The Internet might be more important to them in general and consequently they are willing to accept a higher price even for BE Internet.

Participants with a job that is more likely to result in a higher monthly income also demand a higher monetary compensation for BE. This effect is significant on the 5 percent level for relative quality provision, but only on the 7.5 percent level for dedicated quality provision.

Participants with such jobs are more likely to have higher opportunity costs and are presumably more likely to fear restrictions than other participants. Somewhat puzzling is the negative effect of usage intensity on monetary compensation that can be observed under relative quality provision. The only reason for this result could be a proportion of participants that would perhaps use only a handful of services in the Internet, but those very frequently (e. g. , e-mail). *Sensitivity* has no significant effect and therefore it is even more likely that the effect is caused by usage intensity of insensitive services. These users would perhaps decide for BE on their own and are happy to pay less compared to NN. With this intention in mind, it would not be surprising to find that these users are willing to pay comparably more for BE service.

Table 4.11.: Estimates - Compensation BE

	Relative QoS			Dedicated QoS		
	Estimate	St. Estimate	P	Estimate	St. Estimate	P
Ranking	-0.121	-0.013	0.506	**-0.513**	**-0.066**	***
Complexity	-0.107	-0.017	0.696	-0.262	-0.042	0.329
Transparency	-0.394	-0.045	0.353	-0.712	-0.082	0.088
Care	**1.129**	**0.134**	*	**1.230**	**0.149**	**
Professionalism	**-1.516**	**-0.111**	*	-1.256	-0.094	#
Equity	-0.731	-0.086	0.137	-0.458	-0.055	0.343
Equality	0.134	0.020	0.721	0.556	0.084	0.129
Knowledge	**-0.787**	**-0.108**	**	**-0.592**	**-0.083**	*
Age	**0.102**	**0.135**	***	**0.107**	**0.144**	***
Gender	-0.190	-0.015	0.666	-0.157	-0.013	0.717
Sensitivity	1.168	0.034	0.335	0.879	0.026	0.459
Satisfaction	-0.148	-0.024	0.485	-0.159	-0.026	0.443
Degradation	-0.468	-0.025	0.449	0.288	0.016	0.634
Problem	0.415	0.019	0.555	0.058	0.003	0.932
Intensity	**-0.160**	**-0.088**	*	-0.111	-0.062	0.119
Bill Fixed	0.009	0.027	0.423	0	0.001	0.983
Bill Mobile	0.015	0.047	0.154	0.018	0.058	0.078
Time	**0.170**	**0.097**	**	**0.117**	**0.068**	*
Household	0.137	0.026	0.413	0.044	0.008	0.788
Profession	**0.341**	**0.075**	*	0.281	0.063	#

# p<0.075, * p<0.05, ** p< 0.01, *** p < 0.001

4.8.5. Status Quo

Table 4.12 shows the estimated regression weights of all variables that are assumed to influence the presumption to be artificially degraded by the ISP. *Sensitivity* makes it indeed more

Table 4.12.: Estimates - Degradation

	Degradation		
	Estimate	St. Estimate	P
Problem	**0.252**	**0.217**	***
Sensitivity	**0.162**	**0.088**	*
Gender	-0.015	-0.023	0.507
Age	0.003	0.066	0.078
Knowledge	0.011	0.029	0.457
Equality	-0.038	-0.104	#
Equity	-0.023	-0.050	0.350
Care	**0.063**	**0.141**	**
Transparency	-0.023	-0.049	0.283
Intensity	-0.001	-0.01	0.796
Bill Fixed	**-0.002**	**-0.086**	**
Bill Mobile	**0.002**	**0.140**	***
Professionalism	-0.049	-0.067	0.201
Time	0.004	0.042	0.207
Household	0.000	0.001	0.976
Profession	0.008	0.031	0.365

# p<0.075, * p<0.05, ** p< 0.01, *** p < 0.001

likely to assume to be artificially degraded by the ISP. The same holds true if a participant is more likely to attribute technical problems to the ISP. Based on these results one can verify Hypothesis 4.4 and 4.5. Furthermore, one can observe a positive effect of *Care* on the presumption to be artificially degraded by the ISP. Participants who perceive *Care* as an important aspect in the relation to the ISP are more likely to be suspicious about the delivered service quality under the status quo. The highly significant effect of *Bill Mobile* is positive, whereas the significant effect of *Bill Fixed* is negative. This finding supports for purported difference between the behavior of ISPs in mobile and fixed networks. Heavy mobile users are significantly more likely to presume to be artificially degraded, because they are already more likely to have experienced a form of degradation. Mobile network operators today enforce numerous restrictions (e. g. , VoIP throttling, data caps and speed limits) and IUs obviously adapt their expectations accordingly. These effects have an indirect implication on the satisfaction with the status quo. Table 4.13 shows the estimated regression weights of all variables that are assumed to influence the perception of the status quo. Both, the attribution of problems to the

Table 4.13.: Estimates - Satisfaction

	Satisfaction		
	Estimate	St. Estimate	P
Degradation	**-0.330**	**-0.109**	**
Problem	**-0.495**	**-0.141**	***
Usage sensitivity	-0.238	-0.043	0.270
Gender	-0.004	-0.002	0.955
Age	0.002	0.016	0.699
Knowledge	0.020	0.017	0.696
Equality	-0.013	-0.011	0.850
Equity	-0.039	-0.029	0.635
Appreciation	0.038	0.028	0.624
Transparency	0.067	0.047	0.355
Usage	0.005	0.017	0.700
Bill Fixed	-0.001	-0.026	0.476
Bill Mobile	-0.001	-0.028	0.437
Professionalism	-0.039	-0.018	0.766
Time	0.018	0.063	0.089
Household	-0.020	-0.024	0.496
Profession	-0.001	-0.002	0.968

# p<0.075, * p<0.05, ** p< 0.01, *** p < 0.001

ISP as well as the presumption to be artificially degraded by the ISP have a significant negative effect on the satisfaction with the status quo. This findings confirm Hypotheses 4.6 and 4.7. ISPs therefore indeed reduce their consumers satisfaction with non-neutral network management procedures. They also face drawbacks from nontransparent management of network congestion. IUs are more likely to presume to be degraded if they attribute technical problems to the ISP. Moreover, Table A.6 in the appendix reveals that the attribution of technical problems to the ISP has an additional significant indirect effect on the satisfaction with the status quo (partial mediation). Therefore, ISPs would do better if they communicate their network status more detailed to customers and prevent the attribution of problems to their business.

4.8.6. Profitability of Internet User tiering

The question remains if it is profitable for ISPs to introduce IU tiering to their customer base. Recall that the preference for IU tiering has a positive significant effect on the WTP for IU tiering. Even if the effect is not huge, it makes sense to take into consideration implementing the most preferred quality provision mechanism to maximize the positive effect on the WTP for priority. Figure 4.9 shows the distribution of the ranking of each quality provision mechanism. The bars indicate how often a network regime has been ranked according to the values presented in Table 4.1. The figure reveals that dedicated quality provision is preferred over the relative quality provision mechanism. The figure reveals that more than one third of all respondents would favor IU tiering with dedicated quality provision over all other alternatives. This means 72.26 percent of all respondents deciding for IU tiering as their most favored network regime would prefer the dedicated quality provision mechanism. This result is strongly in favor of a dedicated quality provision mechanism if the transition to an IU tiering system is planned.

To further analyze the WTP of participants in the survey the difference between the premium for priority and the necessary compensation for BE has been calculated for each respondent in both network regimes. For the reminder of this thesis this difference will be called the *delta*. This new variable is useful to classify participants. A positive delta indicates that a participant has a higher valuation for the anticipated additional quality in the priority lane in relation to the expected disutility of beeing in the BE lane. Conversely, a negative *delta* indicates that a participant expects a higher disutility from being in the BE lane compared to the expected gain from additional quality in the priority lane. The group of participants with a negative delta can be further divided into users who boycott IU tiering and those who don't. Boycott means that

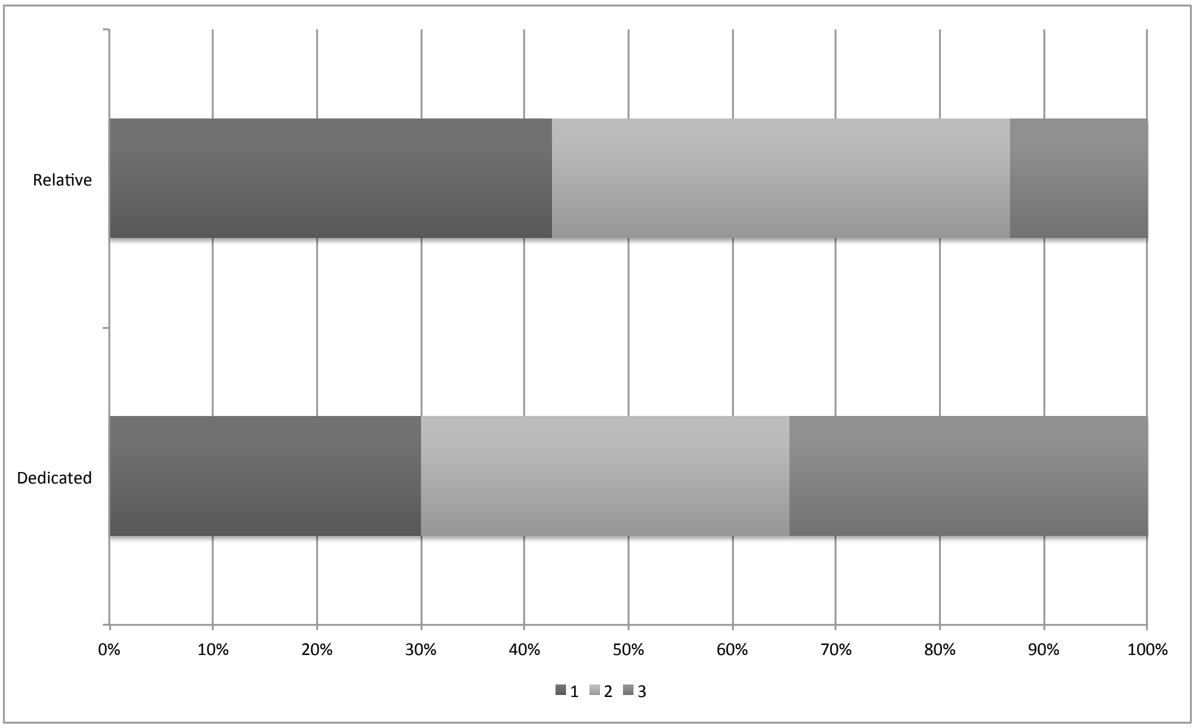

Figure 4.9.: Preference quality provision mechanism

users indicate zero additional WTP for priority and maximum necessary compensation for BE. This means those IUs would accept BE Internet access only for a price of zero and would pay for priority Internet access only the same price as under NN. Otherwise, they implicitly claim to stay out of the market and to buy no Internet access at all. Figure 4.10 shows the distribution of these three distinct groups. The figure reveals, that the share of the boycott group is comparably high with 17-18 percent.

This effect can also be seen in Figure 4.11, which shows the aggregated demand curves for priority and BE under IU tiering. To understand this figure, recall that any customer provides two times a tuple of prices (priority and BE).

It is obvious, that even a monopolistic ISP has no incentive to deviate from the status quo, because any price change from the anchored price under NN (20 Euros) would inevitably drive so many customers out of the market that additional revenues from priority can not

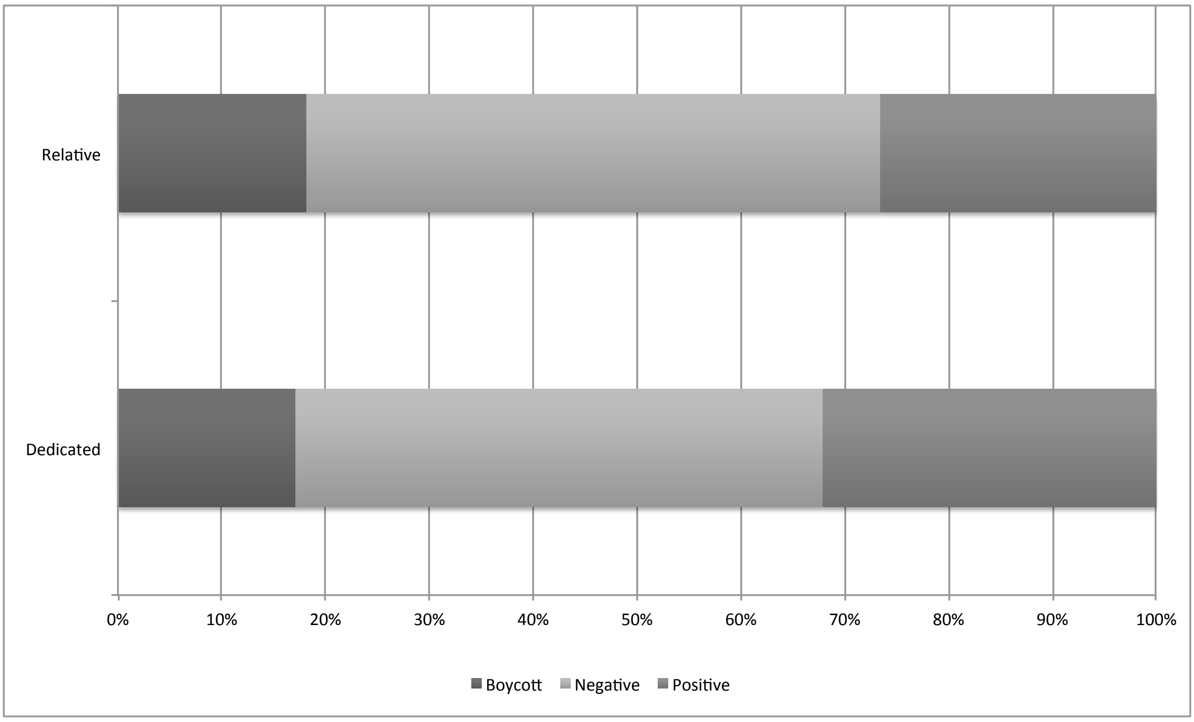

Figure 4.10.: Delta WTP

compensate for the losses.[4] Any deviation from NN would drive the whole boycott group out of the market.

To better understand this effect consider an example. Imagine a monopolistic ISP would calculate its priority price based on the premium that participants report to be willing to pay for priority and not based on the total price for priority access. Based on this assumption and the aggregated demand shown in Figure 4.11, the ISP could calculate the revenue maximizing premium for priority access. Therefore he would charge 9.9 euros extra for priority access under dedicated quality provision and 4.9 euros for priority access under relative quality provision. This would maximize his revenues from priority premiums. As a result of this priority access prices he could expect in the best case that 29.95 percent of the participants would be willing to buy priority access under dedicated quality provision and 49.33 percent of the participants under relative quality provision. Based on the residual demand for BE (each customer can only buy once) the revenue maximizing price for BE service would be 9.9 euros

[4]Under the assumption, that 20 euros is the optimal monopoly price under NN.

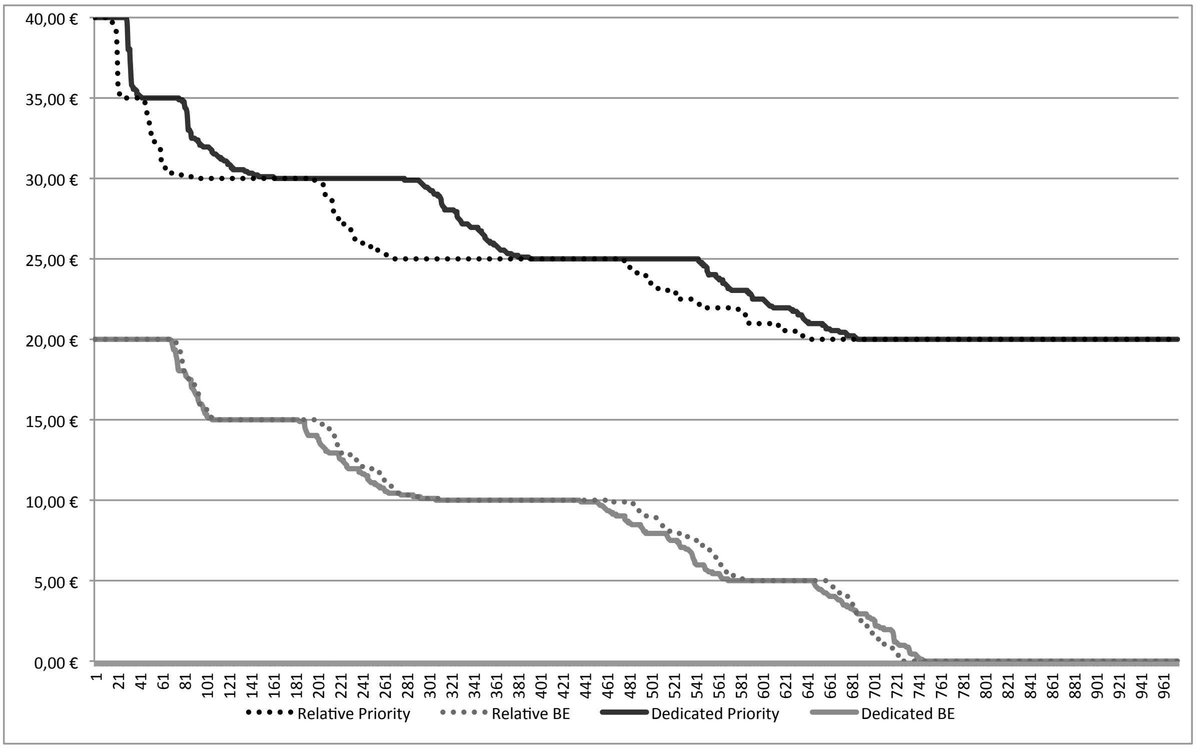

Figure 4.11.: Aggregated demand

in both network regimes. As a result 27.38 percent of the participants would buy BE access under dedicated quality provision and 18.67 percent under relative quality provision. The total revenue under dedicated quality provision would be 58.33 percent of the revenue under NN and under relative quality provision 70.66 percent of the revenue under NN.

Now consider another ISP that decides based on the data, but prices priority and BE services according to the average WTP of its customers. To this end consider the following formula, in which α denotes the share of consumers under IU tiering buying prioriy and WTP represents the average willingness to pay for priority and BE, respectively, assessed independently for each quality provision mechanism.

$$\alpha \frac{\sum_{i=1}^{n} WTP^{P}}{n} + (1-\alpha)\frac{\sum_{i=1}^{n} WTP^{BE}}{n} > 20$$

In order to establish an IU tiering regime with higher revenues compared to NN the ISP has to earn more money with the average customer compared to NN (i. e., more than 20 euros). The average willingness to pay under IU tiering with relative quality provision for priority access is 24.34 euros and for BE access 7.96 euros. Conversely, under IU tiering with dedicated quality provision the average willingness to pay for priority access is 25.47 euros and for BE access 7.80 Euro. In order to increase the revenues compared to NN the ISP would have to convince at least 73.5 percent of the existing customer base to buy priority under an IU tiering regime with relative quality provision and 69 percent under IU tiering with dedicated quality provision. However, given these prices and the data at hand, the necessary share of priority buyers could not be achieved and IU tiering would not be profitable.

Limitations

A limitation w. r. t. Research Question 5 is the sample of users that already pays for Internet access services. BE access for a lower price compared to NN could activate new users to buy Internet access and encourage them to enter the market. With the data at hand, one cannot make any predictions about how many additional users the introduction of a cheaper BE class could attract (if any). This fact relativizes the findings above. With the introduction of service classes in Internet access, the ISP wants to create a separating equilibrium in the market between users with a high valuation for quality and users with a low valuation of quality. However, the surveyed data does not provide information about the participants' decision between priority and BE access if they are in an IU tiering scenario. Therefore, one cannot directly infer what a user would choose if he or she faces a tuple of priority and BE prices that are below his or her WTP for the respective service class. This information is important, because otherwise one cannot make any prediction about how many users would opt for the cheaper service and therefore cannibalize the priority service.

Another important aspect that cannot be further evaluated by the data of this survey are the costs of QoS provisioning. It has been argued in Chapter 4 that many network components today are already capable of QoS, and therefore hardware costs are presumably low. However ISPs have to install new billing systems and have to make the provision of priority transparent to the consumer. On this end investments into the accounting and billing infrastructure might be necessary, but those costs are likely to be fixed and not variable.

In sum it is very doubtful that IU tiering with two classes of services could be a profitable strategy for ISPs. The most likely scenario of successful IU tiering might by the activation of people who currently do not use the Internet because they perceive it as too costly. With the data at hand only such a demand expansion effect could lead to higher profits.

4.9. Implications

Fairness plays an important role in the perception of IU tiering. By contrast, the status quo has no direct influence on the acceptance of new network regimes and the WTP. The presumption about artificial quality degradation has no negative effect on those decisions. However, this factor has an direct effect on the satisfaction with the status quo. Violations of NN (e. g. , to reduce costs) have already a negative impact on customer satisfaction today and consequently on the business of ISPs too. The model suggests further that more technical problems that are attributed to the ISP have an impact on the presumption to be degraded by the ISP. Unintentional restrictions with the Internet access service could be misinterpreted as intentional degradation, or increase the awareness of customers for quality issues.
Consumers who that think *Care* is an important aspect of the relation to the ISP are more likely to assume to be degraded intentionally. One could assume that those customers are more suspicious.

Relative and dedicated quality provision mechanisms are perceived completely different. Relative quality provision mechanisms are less likely preferred if users are afraid of the effort and the complexity that may come along with the introduction of differentiated services. In

addition, it is likely that users perceive relative prioritization as a violation of the *Equality* principle. However, relative priority also suggests the impression of personal and individual treatment. This could explain why *Care* has a positive effect on the preference for relative quality.
In reverse, dedicated quality mechanisms appeal to consumers who have a preference for *Transparency*, and are therefore more likely to be preferred. Moreover, dedicated quality mechanisms are not likely to be less preferred if *Equality* is important to consumers. Quite the contrary is true. *Equity* fairness is associated with this network regime and therefore it appeals to consumers that perceive this concept as suitable for Internet access services.
The WTP for priority under IU tiering shows a quite homogeneous picture. As expected, *Equity* is a main driver for the WTP, but also heavy users are willing to pay more. By contrast, the monetary compensation is mostly influenced by trust in the relation to the ISP. If BE is perceived as a negative event from the customer's point of view, customers with a high personal attachment to the ISP (*Care*) have to be compensated more, while users that rely on the business aspects of the relation to the ISP (*Professionalism*) have to be compensated less.
Furthermore, older people have to be compensated more under BE, which might be due to the fact that younger participants have a higher valuation of the Internet, even with BE service. Users with a better understanding of technical matters are more likely to demand a lower monetary compensation. Perhaps these users do not tend to overestimate the negative adverse effects of prioritization on the BE class.

The profitability of the introduction of IU tiering is highly doubtful (at least in the form presented in this study). A large proportion of the participants decide for NN as their most preferred network regime (52.31 percent) and many users boycott through their indication of a non-existent additional WTP for priority and the rejection of BE service for a positive price. Note that these results were found despite the fact that the emotionally charged term NN was avoided throughout the whole survey.

This could be overcome by clever marketing efforts. Imagine the ISP offers three classes of service (instead of two) and presents the medium priority class as the 'neutral' standard access product for e. g. , 20 euros. Priority access would be priced as a premium option on top. BE access could be marketed as a monthly price discount. This presentation would reduce the impression of a 'two-class' world and still deliver a neutral reference point for existing customers.
Nevertheless, dedicated quality provision shows many favorable characteristics. It triggers the desired aspect of distributive justice and avoids the impression to treat users unequally. The network regime ranking of dedicated quality provision seems not to be influenced by interpersonal justice, which could reduce negative brand effects (i. e. , the preference may not crucially hinge on the reputation of the ISP). Therefore this variant is recommended for an introduction of IU tiering and should be explicitly mentioned in the communication with the potential customers.

Further research is needed to make predictions about potential cannibalization of the priority services classes through BE and what demand expansion effects can be expected from cheaper BE products. It would be also interesting to include the level of provided quality (compared to NN) as a variable. The estimates for the WTP under priority and BE in this study do not advertise any guarantees w. r. t. the delivered quality. To the respondent it is not fully clear how much quality gain or loss he has to expect, even if the animations in the survey were explicitly designed to show an equal gain-loss ratio compared to NN. Conversely, the expectation about the delivered quality is at least partially included in the indicated WTP for priority and BE. Providing respondents with real life examples (e. g. , a live video embedded in the survey, delivered with priority connection or BE connection quality) could lead to more extensive predictions.

Chapter 5.

Neutrality in the Internet Ecosystem

"[...] you can't regulate technology. What you have to do is regulate an industry."

Ivan Seidenberg (CEO of Verizon Communications) (Murray, 2010)

THE notion of neutrality in Internet access services has already flashed over to other parts of the Internet ecosystem. This outlook presents directions of how the topic might evolve in the future. The chapter is based on the paper by Krämer, Wiewiorra, and Weinhardt (2012). Each section provides the main arguments for a neutrality principle in the respective domain and discusses if the principle of neutrality could be legitimately applied in other layers of the Internet Ecosystem as well. Figure 5.1 shows in what ways the principle of neutrality could be applied up and down the Internet value chain.

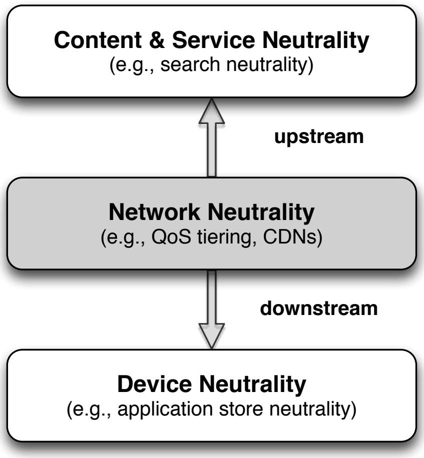

Figure 5.1.: Neutrality in other layers of the Internet ecosystem

5.1. Net Neutrality and Content Distribution Networks

Section 1.4.4 already gave a brief introduction into the business model and the technology of CDN. Figure 1.2 illustrates the possible substitution between the service of an CDN and QoS tiering w. r. t. data transmission quality in the Internet. Both technologies are capable to deliver content or services with a desired level of Quality of Experience to the customers.

Proponents of NN argue that CDNs are no violation of NN, since they deliver data without violating the BE principle, although they do not use the same routes as other traffic. In addition, proponents argue that there is sufficient competition between traditional CDNs. Consequently they can not be seen as gatekeepers to data transmission quality in the Internet.

Opponents (e. g., Yoo, 2005) argue, that the mere existence of CDNs already introduces an entry barrier for new and financially weak CSPs. Many of the existing NNN concerns are therefore not inherent to neutrality, but inherent to market power. This leads to the question if the market for CDNs is indeed sufficiently competitive. Akamai, the biggest and oldest CDN, currently has more than 60 percent market share (Rice, 2010). Clearly, this would be considered a dominant position by any regulatory authority. Therefore, large CDNs should be considered as gatekeepers in the Internet economy.

In the recent years telecommunications providers and other eyeball ISPs have begun to invest heavily in CDN technology and offer content distribution products in their own networks. Local access providers often cooperate with bigger CDN providers and they market these products as upgrades to the basic operational features of their distribution networks (e. g., Rayburn, 2009). Figure 5.2 shows the inherent problem that leads to these cooperations. Traditional CDNs are only able to deliver data to the ‘front door’ of an eyball ISP network because they have no direct control over the internal network infrastructure of ISPs. They can provide end-to-end network quality only in the backbone of the Internet, whereas eyeball ISPs offer only end-to-end network quality in their own local access networks. Both parties are, consequently, unable to offer end-to-end connection quality (see Figure 5.2). One could rather argue, that the services of both sides are complements to each other.

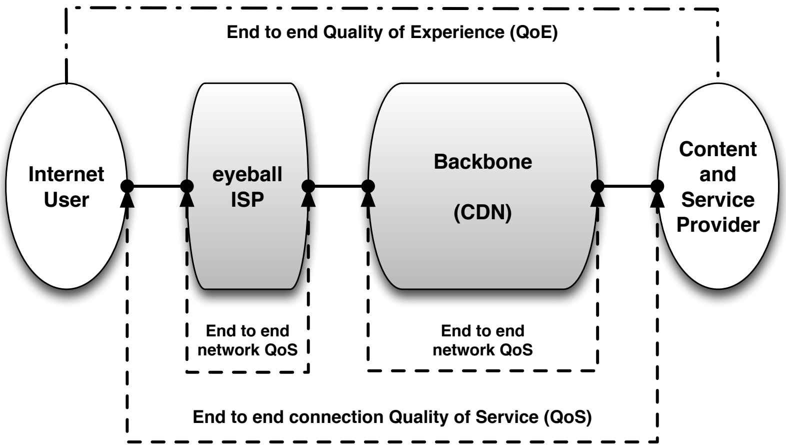

Figure 5.2.: End-to-end service quality (inspired by Statovci-Halimi and Franzl, 2011)

Nevertheless, also under this scenario ISPs use their market power to provide a costly product, giving only paying CSPs a competitive advantage in their own network. It is obvious that even if the technology is different from QoS, the possible negative implications could be similar to the ones discussed in Section 2.2.1. One could argue, that the introduction of CDN technology in eyeball ISP networks is only a technological circumvention of possible NN ruling. Even if data is transported based on the BE principle, market power over the infrastructure allows ISPs to grant quality improvements in their networks to financially strong CSPs exclusively.

5.2. Device Neutrality

Renda (2010) points out that the Internet ecosystem in all it's parts is affected by vertical competition. Players up and down the Internet value chain exert pressure on the network layer by taking over revenue sources previously exploited by ISPs.

In recent years, with the tremendous success of smart phones and tablet computers, mobile device manufacturers (e. g. , Apple) and owners of mobile operating systems (e. g. , Google) are exerting more and more control over the content and services that are consumed on this class of devices. In the days when predominantly feature phones were sold, (mobile) network operators were in control of what software and services were preinstalled one the subsidized devices. On these devices there was no customizable operating system installed, and customers were stuck with whatever software or services were supported by their providers. This power structure has changed dramatically. Nowadays mobile ISP have almost lost complete control over the software on mobile devices. Their own affiliated services (Short Message Service (SMS), Multi Media Message (MMS), video telephony etc.) are becoming obsolete, because their functionality is replaced by applications that rely only on Internet connectivity and not on special network embedded service protocols. Additionally, device manufacturers are opposing wishes from network owners to control the technological capabilities of their handsets. Hahn et al. (2007, p.424) names the following restrictions that network operators try to impose on subsidized devices, sold together with mobile phone contracts.

- “Require handset be sold by operator (or agent of operator)

- Prevent use of handset on rival's network
- Require manufacturers to remove or limit call timers
- Disable certain Bluetooth functionality
- Disable Wi-Fi
- Must use additional service to transfer photos"

In this respect Apple changed the mobile industry fundamentally. The company entered the market with a highly desired product, the iPhone, and offered to the mobile operators a trade-off in the aforementioned dimensions. They agreed to multi year exclusive deals for the iPhone with one single operator per country, giving this certain provider the possibility to extract a (local) monopoly rent on the contracts sold with the phone. Apple also agreed to limit the usage of the handset to this single network. In exchange they claimed full control over the end user experience and software, including all wireless functionalities. After having successfully established a significant market presence they let all exclusive deals expire.

This paradigm shift resulted in a gradual decline of value added services by mobile operators. The elimination of these traditional revenue sources is often viewed as the main driver for the desire to introduce QoS in mobile networks (Wu, 2007; Wallsten, 2007; Hahn et al., 2007). After coining the term Net Neutrality in 2003, Wu extended the concept to the mobile industry by subsuming the related problems with the term 'Wireless Net Neutrality'. In the respective paper Wu (2007) surveys the practices of mobile network operators to exert as much control over their networks as possible (see list above). He calls for regulation that would impose the right to connect any non-harmful device to the network. He also calls for regulation of subsidization of devices and strict rules to prevent 'crippled products', like phones with disabled Bluetooth functionality.

Wallsten (2007) responds to the paper of Wu in a short note. He thinks that many of Wu's arguments are flawed because there is enough competition in the wireless industry, and he lists many counterexamples.

Hahn, Litan, and Singer (2007) call for an moderate approach w. r. t. to wireless NN. They conclude that regulatory intervention is only necessary if there is a definite market failure and if the proposed interventions lead to a better outcome as the status quo. This recommendation, however, seems tautological.

5.3. Content and Service Neutrality

A similar line of argumentation may hold true upstream of the Internet value chain as well. Also on the content and service layer of the Internet ecosystem one can observe a concentration of market power and the emergance of new gatekeepers. One prominent example is Google. Google's search engine is the entry point to the Internet for many IU around the world. The search engine is the first place to go, when looking for information in the WWW. However, no one really knows or can control how the Google search algorithm decides which results are displayed on top. Google only publishes the set of measures that are taken into account, but there is no information on how a single measure is weighted in the final 'page rank'.

Odlyzko (2009) argues that even Sergey Brin and Larry Page, the founders of Google, once thought, that advertisement funded search engines have an incentive to bias search results in favor of their paying advertisers. The problem becomes even more evident if one considers other well known gatekeepers in the modern Internet ecosystem. Social network providers (e. g. , Facebook, Google+) own the information on the so-called 'social graph' (the aggregate information about all links between members of the social network, as well as the related personal information on each member). With this information search engines can even more personalize search results based on personal preferences, social affiliation and the user's browsing history. Grimmelmann (2010, p.438) presents the following eight "[...]search-neutrality principles:

1. *Equality:* Search engines shouldn't differentiate at all among websites.

2. *Objectivity:* There are correct search results and incorrect ones, so search engines should return only the correct ones.

3. *Bias:* Search engines should not distort the information landscape.

4. *Traffic:* Websites that depend on a flow of visitors shouldn't be cut off by search engines.

5. *Relevance:* Search engines should maximize users' satisfaction with search results.

6. *Self-interest:* Search engines shouldn't trade on their own account.

7. *Transparency:* Search engines should disclose the algorithms they use to rank web pages.

8. *Manipulation:* Search engines should rank sites only according to general rules, rather than promoting and demoting sites on an individual basis."

Grimmelmann states that a search engine, that treats all websites equally can not distinguish between relevant and irrelevant information. A search engine should not decide what is correct information because this is in the eye of the user. Each search term may lead to ambiguous results, because the understanding of the term may differ between users. This is also the reason why personalization of search (i. e., unequal treatment) may be the best solution to narrow down the potential meaning of search terms.

In his view the argument of search engines being biased because they favor big traffic intensive websites over small ones is flawed, too. For an algorithm 'relevance' usually means 'information that matters' and information that already matters to a lot of people is likely to be also more important to another user (and not a small website that no one visits).

Furthermore, he states that there is no reason to establish a right for constant traffic. In his view, the exact opposite should be done, because users should not be directed to a site, if it is irrelevant to them. This is the natural free market competition of new content and services, that finally may lead to further innovations. This is one of the underlying arguments of NN and would even be perverted under this principle.

Transparency, in contrast to regulation of Internet access, is no viable option in the search business. The search algorithm itself and the exact composition of relevant parameters used to generate search results are the core business secret of the search engine company. Enforcing transparency would make the service easily duplicable and consequently could kill the business of search engines.

In summary Grimmelmann comes to the conclusion that all of these principles are unusable for any possible form of regulation in the market for Internet search. However, he also comes to a conclusion that is similar to the decision support guideline for NN regulation proposed in Section 2.4: "Just because search neutrality is incoherent, it doesn't follow that search engines deserve a free pass under antitrust, intellectual property, privacy, or other well-established bodies of law. Nor is search-specific legal oversight out of the question. Search engines are capable of doing dastardly things[...]" (from Grimmelmann, 2010, p.438)

Chapter 6.

Conclusions and Future Research

"The Internet needs a makeover. Unfortunately, congressional initiatives aimed at preserving the best of the old Internet threaten to stifle the emergence of the new one."

David Farber and Michael Katz (Farber and Katz, 2007)

THIS work on the economics of NN and the effect of priority pricing in access networks extends the existing NN literature in several ways.

To structure the different approaches, arguments and assumptions about NN and the potential threats of deviations from the status quo a **NNN framework** has been proposed. The framework is structured along two important dimensions: The network regime, which describes the degree of network management, and the pricing regime, which can be either one- or two-sided. This categorization allows to evaluate NNN on the basis of unique scenarios, each of which is based on an individual set of assumptions. In the political debate about NN different scenarios often get mixed up and consequently illegitimate conclusions are drawn.

Furthermore, all threats and possible regulatory remedies of the proposed NNN scenarios have been consolidated in a **policy decision guideline**. The guideline recommends regulators to first check two fundamental assumptions: First, ISPs face (or will face) a congestion problem in the access and backhaul networks and second, ISPs are not able to cope with the increasing costs for additional capacity, and therefore have to generate additional revenues. Only if these assumptions are found to be valid a deviation from the status quo would be necessary. The proposed decision guideline embeds a NN law as the last resort. Each possible threat of the different NNN scenarios can be counteracted by less severe measures. This is important because the mere identification of potential problems does not necessarily imply that ISPs would exploit their technological power under NNN, especially if they are in competition with each other. Only if those measures are assumed (or proven) to fail, a NN law would be suitable. The policy decision guideline complements the NNN framework and allows politicians and regulators to evaluate their available measures to protect the Internet ecosystem from possible negative effects of network management and network pricing.

The reminder of this work analyzed QoS tiering. The two relevant scenarios which have been identified in the NNN framework are CSP tiering and IU tiering.

Chapter 3 analyzed the scenario in which CSP can pay an additional fee to get preferential access to the customer base of individual ISPs. To this end a **theoretical model** has been presented that extends the existing literature by incorporating the aspect of entry (innovation), endogenous quality differentiation and congestion sensitivity. This approach allows to analyze an important aspect of the CSP tiering scenario in a setting with a monopolistic ISP: The interplay between innovation at the edge of the network and innovation at the core.
It has been found that under the assumption of uniform congestion sensitivity CSP tiering is most likely the welfare superior network regime. This has two main reasons: First, in the long-run CSP tiering fosters the entry of new congestion sensitive services that could not establish a sustainable business under NN, and second, the introduction of QoS distributes the congestion more efficiently between the heterogeneous CSPs.

These results stem from the fact that QoS pricing provides the ISP with additional revenues, which in turn allows for higher network investments. Conversely, this means that compared to NN CSPs are worse off, because the ISP is now able to extract some of their revenues through priority pricing. The remaining CSPs are worse off because they are now suffering from higher congestion. However, under uniform congestion sensitivity, no CSP is driven out of the market.

These results obviously depend on the distribution function of congestion sensitivity. It has been shown with a special non-uniform distribution function, that a higher mass of congestion sensitive CSPs could lead to situations in which NN is the superior network regime. MQSs in the context of CSP tiering have to be used with caution. Especially political goals like quality parity between BE and NN could lead to an excessive investment in capacity.

Chapter 4 analyzed the scenario in which ISPs introduce preferential treatment of data to the IU side of the market. Regulators are less concerned about IU tiering because it represents a one-sided pricing scheme (i. e. , only IU pay) and, therefore, many NNN threats are irrelevant. IU tiering is consequently more likely to be accepted by regulators. Therefore, ISPs are interested in a successful transition to and a sustainable business case under IU tiering.

Real world examples in the context of NN show that consumers can outrage about perceived unfair treatment and can, as a consequence, boycott new business practices. This work extends the existing literature by providing an **empirical model** to analyze the effect of fairness on the acceptance of IU tiering and the WTP to pay for preferential access, as well as the necessary monetary compensation to accept BE service. To collect data an online survey among 1035 representative Internet access customers has been conducted. The data has been used to confirm and evaluate a SEM. The results reveal that fairness indeed plays an important role in the perception of IU tiering, and that relative and dedicated quality provision mechanisms are preferred for fundamentally different reasons. Dedicated quality provision is favored by customers who perceive *Equity* as the suitable allocative fairness concept and indicate that *Transparency* about their Internet access is important to them.

In contrast to that, relative quality provision mechanisms rely on the concept of *Trust*, and customers who prefer personal and individual treatment are more likely to choose this network regime. However, customers who are afraid of the complexity that comes along with IU tiering are less likely to opt for the relative quality regime. In addition, it is at least very probable that customers who perceive *Equality* as the suitable concept of allocative fairness in the context of Internet access service are less likely to opt for IU tiering with a relative quality provision mechanism.

The WTP for priority depends to some extend on the preference for the respective quality provision mechanism. Furthermore, customers who perceive *Equity* as the suitable allocative fairness concept have a higher WTP for priority.

The necessary monetary compensation for BE depends on both aspects of *Trust*. While *Professionalism* makes it less likely that customers demand a higher compensation under BE, customers who prefer individual and personal treatment (*Care*) from their ISP are more likely to demand a higher compensation.

All these findings have several important implications for regulators and ISPs alike.

Implications for ISPs ISPs need to consider that customers do not only care about the mere outcome of preferential treatment, but also take into account the mechanism of how this outcome is achieved. Dedicated quality provision is highly preferred by customers, and this preference is less facilitated by customers' trust in the ISP than by the mechanism itself appearing more transparent. The analysis of the profitability of IU tiering (as presented in the survey) revealed that the additional WTP for priority is most likely not sufficient to establish a sustainable business case, neither under dedicated nor relative quality provision. This leads to the conclusion that the introduction of two service classes may not be suitable to facilitate a successful transition to an IU tiering regime. To this end, at least three service classes might be necessary to provide existing users with a 'neutral' reference case.

A business case for this kind of quality differentiation should consider that ISPs with a stronger brand are more likely to be able to charge higher prices for BE, compared to their competitors. If the brand effect leads to higher trust in the ISP, the necessary compensation for BE could be

lower. The WTP for priority does not depend on trust in the ISP, but customers unambiguously have a higher WTP for priority delivered by a dedicated quality provision mechanism. This result is despite the fact that customers receive the same level of quality under relative and dedicated quality provision. Furthermore, customers have a measurable WTP for priority and three classes of service could help ISPs to absorb this premiums from equity driven customers without intimidating the remaining customers.
Since regulators are very concerned about two-sided pricing, ISPs should be aware that IU tiering could be the only option to generate additional revenues in the future.

Degradation in the status quo does not influence the decision about IU tiering. However, it has been shown that the current level of *Satisfaction* is negatively influenced by potential cost reducing measures of network management. ISPs who use such methods are more likely to face an unsatisfied customer base, who may be more likely to switch to a competitor. Furthermore, it has been shown that the attribution of problems with the Internet connection to the ISP has a positive effect on the presumption to be artificially degraded by the ISP. This calls for more transparency under the status quo. ISPs should adequately inform their customers about technical difficulties and other technical restrictions to prevent unfavorable conclusions about their network management practices, which in turn could negatively affect their customers satisfaction.

Implications for Regulators Regulators face the difficult challenge to correctly evaluate existing and necessary network capacities as well as the resulting level of congestion in the networks. Based on these evaluations they have to decide if the claims of ISPs are legitimate and NNN is necessary. Furthermore, they have to evaluate the distribution of congestion sensitivity among CSPs, because CSP tiering might only be superior to NN if not too many (future) services are congestion sensitive. They further have to weight the possible gains and threats of certain NNN scenarios against each other and, if necessary, implement remedies to resolve potential issues. However, ex-ante NN regulation should only be the last resort if regulators expect or find all other measures to fail. Otherwise NN obligations could hinder investments in the network infrastructure. Faulhaber and Farber (2010) look into the 2008 spectrum auction (700 Mhz) in the US which can be considered a natural experiment in NN regulation, because one spectrum block was attached to NN obligations. The authors find that the attached NN obligations "[...]decreased the value of the spectrum asset by 60%" (Faulhaber and Farber, 2010, p.331).

The implications about CSP tiering are especially relevant for quasi monopolistic Internet access markets like in the US. Further research is needed to make predictions about the possible effect of QoS tiering in competitive market environments. However, it is very likely that the findings of the monopolistic setting carry over to a duopolistic setting, in which CSPs multi-home (i. e., are connected to each ISP individually). Even if regulators come to the conclusion that a deviation from the status quo would not be necessary per se, the question remains why ISPs should be forbidden to manage their networks to generate additional revenues, if QoS tiering is found to be welfare superior to NN. However, in the current debate, ISPs focus predominantly on arguments about impending congestion and the resulting investment problems.

Future Research The research presented in this work could be extended in several ways. Consider first the theoretical model presented in Chapter 3. Section 2.3.1 summarized the paper of Reggiani and Valletti (2012) that already considers an extension of the model to a setting with CSPs with different market power. The question remains how QoS tiering performs in a competitive environment. The discussion exemplified the most likely result under multi-homing. In reality, however, the Internet is an interconnected network of networks, and quality provision depends on the agreements between different ISPs. Therefore, it would be interesting to further analyze the interplay between and the forming of coalitions of different ISPs offering QoS products. Furthermore, the degree of competition between ISPs might depend on the number of service classes. Classic economic literature (e. g., Shaked and Sutton, 1982) suggests that two competitors are able to abate competition through quality differentiation. However, this result may depend on the number of service classes (i. e., qualities).

The research project about IU tiering could be further extended to analyze possible profitable IU tiering business cases. To this end, it is fruitful to find out how users perceive a deviation to more than two service classes. Future studies have to survey explicitly the preference for a certain service class to allow for more sophisticated profitability analysis. Furthermore, future studies should consider a representative sample of the whole society instead of focusing on Internet access customers exclusively. This would allow to estimate a possible demand expansion effect through the introduction of cheaper BE tariffs. As a consequence, a pen and paper questionnaire would have to be used because otherwise the sample would be biased in favor of participants with Internet access. The study presented in this work focused on the prioritization of individual users. However, ISPs could also decide to grant preferential access based on single services or service categories that have to be specified by the customers. The question remains if a segregated approach to priority pricing could lead to higher revenues. This approach to IU tiering could also help to reduce the number of participants implicitly claiming to boycott a new network regime.

Appendix A.

Internet User Tiering

A.1. Survey

1) Please indicate your agreement to the following statements (1=disagree/5=agree).

◯ I am satisfied with my Internet service provider.

◯ I am satisfied with my Internet connection quality.

◯ I am responsible for the contract of my Internet access.

2) Please indicate your average Internet usage time per day

◯ in fixed access networks:

◯ in mobile networks:

3) How much do you pay on average per month for telecommunications services (Internet, telephony, TV)?

◯ in fixed access networks:

◯ in mobile access networks:

Please answer the following questions with respect to your Internet consumption in fixed line access networks. The following section presents different concepts to handle data traffic in Internet access networks. To visualize the different concepts please imagine the Internet as a highway and the data packets as cars driving on this highway. As on reals roads, a high volume of traffic can cause congestion.

4) Please compare concept A and B with each other. (cf. Figure A.1)

Highlighted (colored) cars represent data packets that arrive simultaneously at a congested area. The blue data packet is treated like all other data packets on the road. The golden euro symbol represents the eligibility to bypass a congested area. The green car has this eligibility, whereas the red car has not.

◯ **Concept A:** All data packets are treated equally in the congested area.

◯ **Concept B:** For an additional fee data packets can bypass a congested area. Other data packets that have not bought this eligibility have to wait in such a situation until those packets passed by.

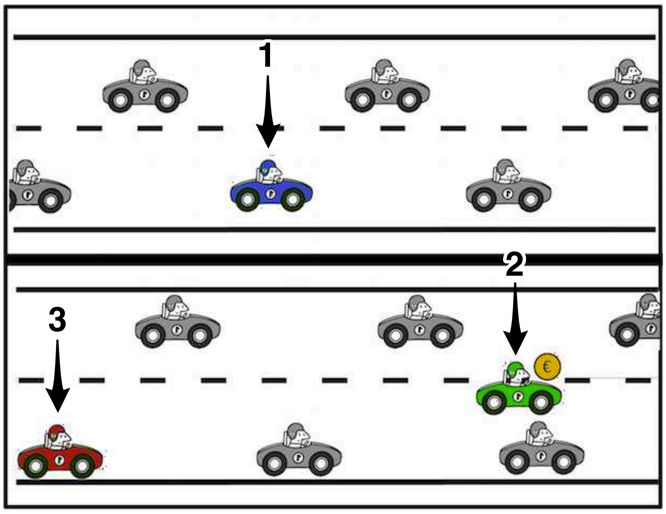

Figure A.1.: NN vs. IU tiering relative quality

5) What is the characteristic of the blue data packet?

◯ The data packet is treated like all other data packets.

◯ The data packet is eligible to bypass a congested area.

◯ The data packet has to wait in a congestion situation until other data packets passed by.

6) What is the characteristic of the gree data packet?

◯ The data packet is treated like all other data packets.

◯ The data packet is eligible to bypass a congested area.

◯ The data packet has to wait in a congestion situation until other data packets passed by.

7) What is the characteristic of the red data packet?

◯ The data packet is treated like all other data packets.

◯ The data packet is eligible to bypass a congested area.

◯ The data packet has to wait in a congestion situation until other data packets passed by.

8) Pleas imagine Internet access according to concept A (blue data packet) costs 20 euros per month. Ho much would you be willing to pay for Internet access with priority according to concept B (green data packet)?

◯ Slider: 20-40 euros.

9) Pleas image Internet access according to concept A (blue data packet) costs 20 euros per month. Ho much would you be willing to pay for Internet access without priority according to concept B (red data packet)?

◯ Slider: 0-20 euros.

10) Please compare concept A and C with each other. (cf. Figure A.2)

Highlighted (colored) cars represent data packets that arrive simultaneously at a congested area. The blue data packet is treated like all other data packets on the road. The golden euro symbol represents the eligibility to bypass a congested area. The green car has this eligibility, whereas the red car has not.

◯ **Concept A:** All data packets are treated equally in the congested area.

◯ **Concept C:** One lane of the road can only be used for an additional fee. As a result the traffic on this lane and the danger of congestion is reduced.

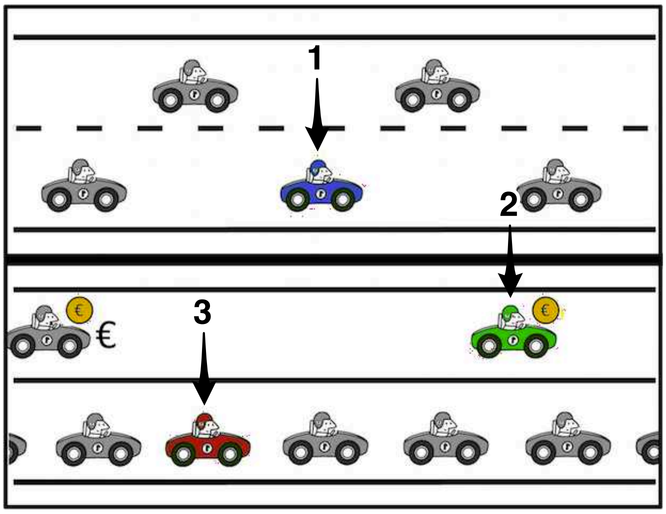

Figure A.2.: NN vs. IU tiering dedicated quality

11) Pleas imagine Internet access according to concept A (blue data packet) costs 20 euros per month. Ho much would you be willing to pay for Internet access with priority according to concept B (green data packet)?

◯ Slider: 20-40 euros.

12) Pleas imagine Internet access according to concept A (blue data packet) costs 20 euros per month. Ho much would you be willing to pay for Internet access without priority according to concept B (red data packet)?

◯ Slider: 0-20 euros.

13) Please compare concept B and C with each other. (cf. Figure A.3)

Highlighted (colored) cars represent data packets that arrive simultaneously at a congested area. The golden euro symbol represents the eligibility to bypass a congested area. The green car has this eligibility, while the red car has not.

◯ **Concept B:** For an additional fee data packets can bypass a congested area. Other data packets that have not bought this eligibility have to wait in such a situation until those packets passed by.

◯ **Concept C:** One lane of the road can only be used for an additional fee. As a result the traffic on the lane and the danger of congestion is reduced.

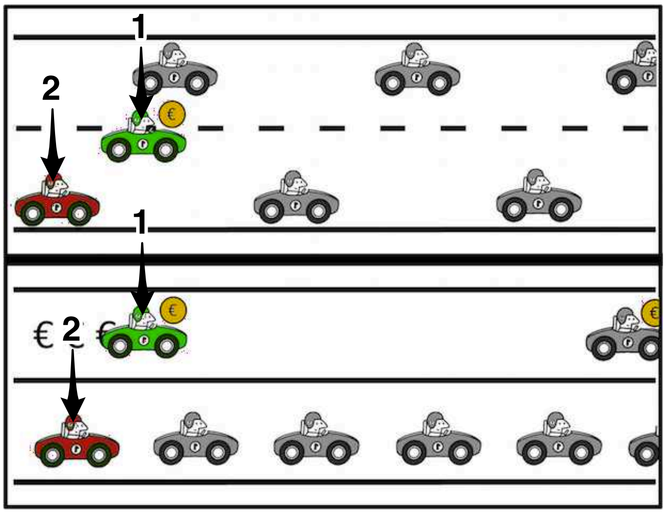

Figure A.3.: IU tiering relative quality vs. IU tiering dedicated quality

14) Please indicate how often you use services or websites from the following category: Internet telephony

◯ never

◯ rarely

◯ occasionally

◯ often

◯ regularly

15) If you have problems accessing services or websites from this category. What is in your opinion the most likely reason?

◯ My Internet Service Provider.

◯ The provider of the service or website.

◯ Other users, trying to use this service or website at the same time.

◯ My technical equipment. (e. g. , Wireless Fidelity (WiFi), computer etc.)

16) Do you think that your Internet Service Provider at least once blocked or degraded services or websites from this category?

◯ Yes.

◯ No.

17) Please indicate how often you use services or websites from the following category: Realtime online gaming

◯ never

◯ rarely

◯ occasionally

◯ often

◯ regularly

18) If you have problems accessing services or websites from this category. What is in your opinion the most likely reason?

◯ My Internet Service Provider.

◯ The provider of the service or website.

◯ Other users, trying to use this service or website at the same time.

◯ My technical equipment. (e. g. , WiFi, computer etc.)

19) Do you think that your Internet Service Provider at least once blocked or degraded services or websites from this category?

◯ Yes.

◯ No.

20) Please indicate how often you use services or websites from the following category: Streaming

◯ never

◯ rarely

◯ occasionally

◯ often

◯ regularly

21) If you have problems accessing services or websites from this category. What is in your opinion the most likely reason?

◯ My Internet Service Provider.

◯ The provider of the service or website.

◯ Other users, trying to use this service or website at the same time.

◯ My technical equipment. (e. g. , WiFi, computer etc.)

22) Do you think that your Internet Service Provider at least once blocked or degraded services or websites from this category?

◯ Yes.

◯ No.

23) Please indicate how often you use services or websites from the following category: E-Mail

◯ never

◯ rarely

◯ occasionally

◯ often

◯ regularly

24) If you have problems accessing services or websites from this category. What is in your opinion the most likely reason?

◯ My Internet Service Provider.

◯ The provider of the service or website.

◯ Other users, trying to use this service or website at the same time.

◯ My technical equipment. (e. g. , WiFi, computer etc.)

25) Do you think that your Internet Service Provider at least once blocked or degraded services or websites from this category?

◯ Yes.

◯ No.

26) Please indicate how often you use services or websites from the following category: Social networking

◯ never

◯ rarely

◯ occasionally

◯ often

◯ regularly

27) If you have problems accessing services or websites from this category. What is in your opinion the most likely reason?

◯ My Internet Service Provider.

◯ The provider of the service or website.

◯ Other users, trying to use this service or website at the same time.

◯ My technical equipment. (e. g. , WiFi, computer etc.)

28) Do you think that your Internet Service Provider at least once blocked or degraded services or websites from this category?

◯ Yes.

◯ No.

29) Please indicate how often you use services or websites from the following category: File sharing

◯ never

◯ rarely

◯ occasionally

◯ often

◯ regularly

30) If you have problems accessing services or websites from this category. What is in your opinion the most likely reason?

◯ My Internet Service Provider.

◯ The provider of the service or website.

◯ Other users, trying to use this service or website at the same time.

◯ My technical equipment. (e. g. , WiFi, computer etc.)

31) Do you think that your Internet Service Provider at least once blocked or degraded services or websites from this category?

◯ Yes.

◯ No.

32) Please indicate your agreement to the following statements (1=disagree/5=agree).

◯ I know a lot about the technical functionality of the Internet.

◯ In my circle of friends and my family I am considered an expert for computers.

◯ I have in-depth knowledge about computer networks and data transmission technologies.

◯ I often need help with computer problems.

33) Please indicate your agreement to the following statements (1=disagree/5=agree).

◯ I think using the Internet is easier, if all services and websites are treated equally.

◯ The usage of the Internet becomes difficult, if I have to consider the option to prioritize data.

◯ I perceive it as clearer, if all services and websites are receive the same network quality.

◯ The option to prioritize data makes the usage of the Internet cumbersome for me.

34) Please indicate your agreement to the following statements (1=disagree/5=agree).

◯ I feel only comfortable using the Internet, if I am acquainted with all details of my Internet tariff and the respective connection quality..

◯ It is important to me to have easy access to all information related to the provided network quality.

◯ It is important to me that all information about the provided network quality are unambiguous and understandable..

◯ I feel uncomfortable using the Internet, if information about my Internet access tariff and the respective connection quality are not verifiable.

◯ I feel uncomfortable, if contractual and technical restrictions of my Internet access service are not revealed to me.

35) Please indicate your agreement to the following statements (1=disagree/5=agree).

◯ Only if every Internet user gets the same network quality I perceive this as fair.

◯ Only uniform transportation of data makes sure, that all Internet users have the same prerequisites.

◯ If network capacity is split up equally between all Internet users, this is fair.

◯ Nobody should be privileged in data transportation in the Internet.

◯ Nobody should be disadvantaged in data transportation in the Internet.

36) Please indicate your agreement to the following statements (1=disagree/5=agree).

◯ It is important to me that my Internet Service Provider is honest.

◯ It is important to me that my Internet Service Provider is reliable.

◯ It is important to me that my Internet Service Provider understands me.

◯ It is important to me that my Internet Service Provider is professional.

◯ It is important to me that my Internet Service Provider takes care of me.

37) Please indicate your agreement to the following statements (1=disagree/5=agree).

◯ It is fair that users who are willing to pay a premium get a prioritized Internet connection.

◯ I think it is unfair, if other users get a prioritized Internet connection, because they are willing to pay a premium.

◯ If a user wants to pay less than other users, it is fair that he gets a lower connection quality.

◯ Prioritization of data for a premium fee is unfair, because other users experience a lower connection quality because of that.

38) Please indicate your gender.

◯ female

◯ male

39) Please indicate your age.

◯ under 18 years

◯ 18 - 29 years

◯ 30 - 39 years

◯ 40 - 49 years

◯ 50 - 59 years

◯ 60 years and older

40) How many people live in your household?

◯ 1 person

◯ 2 persons

◯ 3 persons

◯ 4 persons

◯ 5 persons or more

41) Please indicate your job category.

◯ Apprenticeship

◯ Student

◯ Homemaker

◯ Employee

◯ Worker

◯ Civil servant

◯ Executive

◯ Entrepreneur

◯ Freelancer

◯ Temporary worker

◯ Retired

◯ Seeking work

42) Please allow for a last conclusive question. Did you answer all questions honestly?

◯ Yes

◯ No

A.2. Extensive Model Specification

Note that all fairness and control variables are allowed to covary with each other (indicated by double arrows at the grouping boxes). Furthermore, the error terms of the ranking variables, as well as the error terms of the WTP and compensation variables are allowed to covary with each other. These relations are necessary because the decisions are not independent from each other. For instance, any influence which is not explicitly considered by the model but may cause a higher WTP for priority under relative quality provision, will also partially influence the WTP for priority under dedicated quality provision. The covariances and correlations are reported in Table A.8 and A.9.

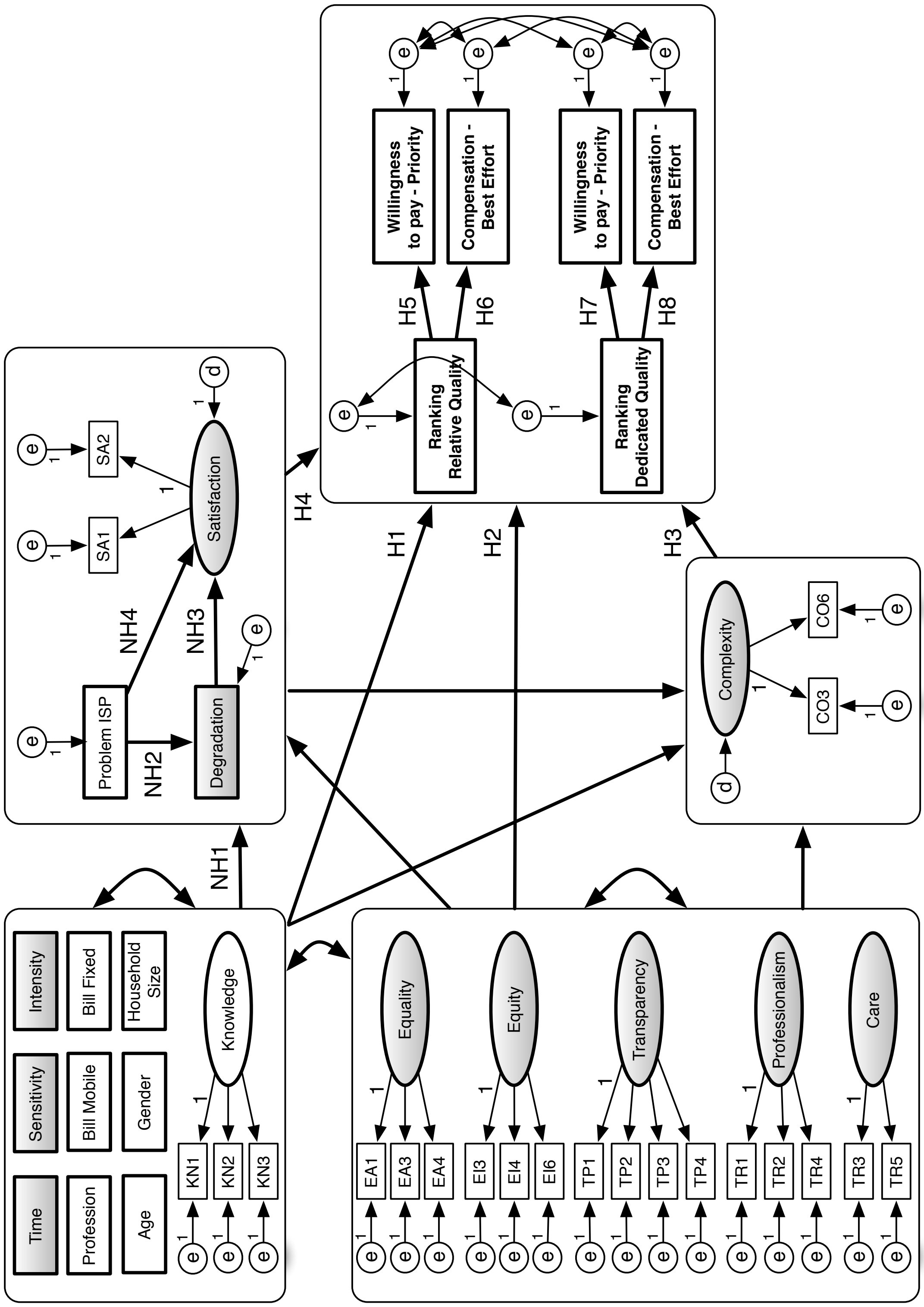

Figure A.4.: Extensive SEM structure

A.3. Model Fit

Table A.1.: Fit - Parsimony

Model	PRATIO	PNFI	PCFI
Default model	,572	,544	,560
Saturated model	,000	,000	,000
Independence model	1,000	,000	,000

Table A.2.: Fit - AIC

Model	AIC	BCC
Default model	1533.238	1567.071
Saturated model	1638.000	1708.150
Independence model	15341.003	15347.684

A.4. Estimates

Table A.3.: Estimates - Problem ISP

	Problem ISP		
	Estimate	St. Estimate	P
Usage sensitivity	-0.051	-0.032	0.375
Gender	0.001	0.002	0.954
Age	0.003	0.076	#
Knowledge	-0.004	-0.013	0.749
Equality	-0.011	-0.036	0.528
Equity	-0.041	-0.105	0.063
Appreciation	0.024	0.061	0.254
Transparency	-0.008	-0.021	0.663
Usage	-0.001	-0.007	0.874
Bill Fixed	0.000	0.017	0.622
Bill Mobile	0.001	0.063	#
Professionalism	-0.043	-0.068	0.216
Time	-0.001	-0.014	0.688
Household	0.004	0.017	0.599
Profession	-0.006	-0.030	0.402

# p<0.075, * p<0.05, ** p< 0.01, *** p < 0.001

Table A.4.: Estimates - Complexity

	Complexity		
	Estimate	St. Estimate	P
Knowledge	**-0.134**	**-0.119**	**
Equality	**0.131**	**0.124**	*
Appreciation	**-0.178**	**-0.136**	**
Transparency	**0.364**	**0.266**	***
Equity	**-0.474**	**-0.358**	***
Age	-0.007	-0.063	0.091
Gender	0.165	0.084	*
Usage sensitivity	-0.063	-0.012	0.735
Satisfaction	-0.024	-0.025	0.466
Usage	0.009	0.031	0.437
Bill Fixed	0.001	0.025	0.439
Bill Mobile	-0.001	-0.016	0.618
Degradation	0.005	0.002	0.957
Problem	-0.159	-0.047	0.142
Professionalism	0.039	0.018	0.729
Time	-0.004	-0.013	0.700
Household	0.031	0.044	0.208
Profession	0.005	0.006	0.855

# p<0.075, * p<0.05, ** p< 0.01, *** p < 0.001

Table A.5.: Estimates - Intercepts

	Relative QoS		Dedicated QoS	
	Estimate	P	Estimate	P
Ranking	**1.129**	***	**1.792**	***
Priority	-0.194	0.856	2.100	0.081
Best Effort	**5.997**	***	**7.352**	***

# p<0.075, * p<0.05, ** p< 0.01, *** p < 0.001

Variables that have no significant indirect effect on at least one of the dependent variables in the model are omitted for clarification.

Table A.6.: Indirect Effects (Part 1)

	Degradation	Satisfaction	Complexity	Ranking Relative	Ranking Dedicated
Equity	-0.010	**0.031** (*)	0.007	**0.047** (**)	-0.020
Knowledge	-0.001	-0.001	0.000	**0.015** (*)	-0.004
Transparency	-0.002	0.012	-0.001	**-0.040** (**)	0.013
Care	0.006	**-0.034** (**)	-0.003	**0.023** (*)	-0.003
Equality	-0.003	0.019	0.001	**-0.016** (*)	0.003
Bill Mobile	0.000	**-0.001** (**)	0.000	0.000	0.000
Age	**0.001** (*)	**-0.002** (**)	0.000	0.001	0.000
Gender	0.000	0.004	0.000	**-0.018** (*)	0.005
Problem ISP	0.000	**-0.083** (***)	0.015	0.035	-0.012
Complexity	0.000	0.000	0.000	0.000	0.000

* p<0.05, ** p< 0.01, *** p < 0.001

Table A.7.: Indirect Effects (Part 2)

	Compensation BE Relative	WTP Priority Relative	Compensation BE Dedicated	WTP Priority Dedicated
Equity	0.036	0.123	0.000	**0.235** (*)
Knowledge	0.003	0.036	0.027	0.056
Transparency	-0.031	-0.129	-0.198	-0.042
Care	-0.023	**0.145** (*)	0.122	0.038
Equality	0.010	-0.088	-0.043	-0.061
Bill Mobile	0.000	0.002	0.000	0.002
Age	0.000	**0.008** (*)	0.002	0.005
Gender	-0.014	-0.017	0.032	-0.131
Problem ISP	-0.019	0.042	0.176	-0.009
Complexity	0.013	**-0.056** (***)	-0.018	0.016

* p<0.05, ** p< 0.01, *** p < 0.001

Variables that do not significantly covary with each other are omitted for clarification.

Table A.8.: Covariances and Correlations (Part 1)

Parameter			Covariance	P	Correlation
Age	<–>	Transparency	0.692	***	0.115
Age	<–>	Intensity	-10.792	***	-0.370
Age	<–>	PANPANF	39.703	***	0.261
Age	<–>	Profession	4.709	***	0.326
Age	<–>	Professionalism	0.366	**	0.094
Age	<–>	Time	-2.842	**	-0.094
Age	<–>	Household Size	0.861	*	0.086
Age	<–>	Equality	0.630	*	0.080
Age	<–>	Care	0.496	*	0.078
Age	<–>	Knowledge	-0.512	*	-0.070
Care	<–>	Equity	-0.092	***	-0.163
Care	<–>	Transparency	0.268	***	0.491
Error RD	<–>	Error RR	-0.207	***	-0.415
Error PD	<–>	Error BD	-8.039	***	-0.285
Error PD	<–>	Error PR	16.508	***	0.813
Error PD	<–>	Error BR	-8.289	***	-0.288
Error BD	<–>	Error BR	29.430	***	0.821
Error BD	<–>	Error PR	-6.307	***	-0.247
Error PR	<–>	Error BR	-6.638	***	-0.255
Equality	<–>	Care	0.165	***	0.233
Equality	<–>	Transparency	0.240	***	0.356
Equality	<–>	Equity	-0.495	***	-.709
Gender	<–>	Knowledge	0.161	***	0.368
Gender	<–>	Intensity	0.227	***	0.130
Gender	<–>	Equity	0.031	*	0.082
Household Size	<–>	Knowledge	0.077	*	0.073
Time	<–>	Knowledge	0.709	***	0.224
Time	<–>	Profession	-0.750	***	0.126
Time	<–>	Care	0.233	*	0.085

* p<0.05, ** p< 0.01, *** p < 0.001

Table A.9.: Covariances and Correlations (Part 2)

Parameter			Covariance	P	Correlation
Bill Fixed	<–>	Bill Mobile	54.041	***	0.149
Bill Fixed	<–>	Household Size	3.444	***	0.158
Bill Fixed	<–>	Profession	4.369	***	0.076
Bill Fixed	<–>	Time	4.424	*	0.067
Bill Mobile	<–>	Profession	2.971	***	0.022
Bill Mobile	<–>	Time	7.807	**	0.108
Bill Mobile	<–>	Care	1.550	**	0.103
Bill Mobile	<–>	Professionalism	0.704	*	0.076
Professionalism	<–>	Equality	0.141	***	0.324
Professionalism	<–>	Care	0.222	***	0.632
Professionalism	<–>	Transparency	0.183	***	0.546
Professionalism	<–>	Equity	-0.075	***	-0.216
Sensitivity	<–>	Time	0.107	***	0.160
Sensitivity	<–>	Knowledge	0.037	***	0.228
Sensitivity	<–>	Gender	0.020	***	0.216
Sensitivity	<–>	Age	-0.176	***	-0.114
Sensitivity	<–>	Intensity	0.272	***	0.424
Sensitivity	<–>	Profession	-0.025	**	-0.054
Sensitivity	<–>	Transparency	0.015	**	0.116
Sensitivity	<–>	Professionalism	0.008	*	0.092
Sensitivity	<–>	Household Size	0.016	*	0.072
Sensitivity	<–>	Care	0.011	*	0.076
Transparency	<–>	Equity	-0.134	***	-0.249
Transparency	<–>	Knowledge	0.072	***	0.114
Intensity	<–>	Transparency	0.368	***	0.147
Intensity	<–>	Knowledge	0.876	***	0.287
Intensity	<–>	Bill Mobile	10.440	***	0.150
Intensity	<–>	Professionalism	.197	***	0.122
Intensity	<–>	Time	3.516	***	0.279
Intensity	<–>	Profession	-0.794	***	-0.111
Intensity	<–>	Equity	0.243	**	0.094
Intensity	<–>	Care	0.194	*	0.074
Intensity	<–>	Bill Fixed	-4.451	*	-0.070

* p<0.05, ** p< 0.01, *** p < 0.001

Legend: Covariances and Correlations

- **Error RR:** Error term *Ranking Relative Quality*
- **Error RD:** Error term *Ranking Dedicated Quality*
- **Error PR:** Error term *Willingness to pay - Priority Relative*
- **Error PD:** Error term *Willingness to pay - Priority Dedicated*
- **Error BR:** Error term *Compensation - Best Effort Relative*
- **Error BD:** Error term *Compensation - Best Effort Dedicated*

Appendix B.

List of Abbreviations

AMOS Analysis of Moment Structures

AVE Average Variance Explained

BE Best Effort

CDN Content Distribution Network

CFA Confirmatory Factor Analysis

CFI Comparative Fit Index

CSP Content and Service Provider

DPI Deep Packet Inspection

DSL Digital Subscriber Line

EFA Exploratory Factor Analysis

EU European Union

FCC Federal Communications Commission

HTML Hyper Text Markup Language

HTTP Hyper Text Transfer Protocol

IFI Incremental fit Index

IISM Institute of Information Systems and Management

INDEX Internet Demand Experiment

IP Internet Protocol

IPTV Internet Protocol Television

ISP Internet Service Provider

ISDN Integrated Services Digital Network

IT Information Technology

IU Internet User

KIT Karlsruhe Institute of Technology

LTE Long Term Evolution

ML Maximum Likelihood

MLE Maximum Likelihood Estimation

MMS Multi Media Message

MQS Minimum Quality Standard

NEP Network Equipment Provider

NFI Normed Fit Index

NN Net Neutrality

NNN Non Net Neutrality

P2P Peer-to-Peer

QoE Quality of Experience

QoS Quality of Service

RMSEA Root Mean Square Residual

SEM Structural Equation Model

SIP Session Internet Protocol

SLA Service Level Agreement

SMS Short Message Service

SPSS Statistical Package for the Social Sciences

TAM Technology Acceptance Model

TCP/IP Transmission Control Protocol / Internet Protocol

TLI Tucker Lewis Index

TOS Type of Service

US United States

USA United States of America

UTAUT Unified Theory of Acceptance and Use of Technology

VDSL Very High Data Rate Digital Subscriber Line

VoIP Voice over IP

VPN Virtual Private Network

WiFi Wireless Fidelity

WTP Willingness To Pay

WWW World Wide Web

Bibliography

Adams, J. (1965). Inequity in social exchange. *Advances in Experimental Social Psychology 2*, 267–299.

Analysys Mason (2008, September). Final report for the Broadband Stakeholder Group: The costs of deploying fibre-based next-generation broadband infrastructure. Accessed February 16, 2012, http://www.analysysmason.com/PageFiles/5766/Analysys-Mason-final-report-for-BSG-(Sept2008).pdf.

Armstrong, M. (2006). Competition in two-sided markets. *The RAND Journal of Economics 37*(3), 668–691.

Bandyopadhyay, S. and H. K. Cheng (2006). Liquid pricing for digital infrastructure services. *International Journal of Electronic Commerce 10*(4), 47–72.

Bauer, J. M. (2007). Dynamic effects of network neutrality. *Journal of Communication 1*, 531–547.

Bauer, S., D. Clark, and W. Lehr (2009). The evolution of internet congestion. Working Paper, MIT Internet Traffic Analysis Study, MIT. Retrieved from: http://mitas.csail.mit.edu/papers/Bauer_Clark_Lehr_2009.pdf.

Baumgartner, H. and C. Homburg (1996). Applications of structural equation modeling in marketing and consumer research - A review. *International Journal of Research in Marketing 13*, 139–161.

Bentler, P. M. and D. G. Bonett (1989). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin 88*(3), 588–606.

Bollen, K. A. (1989). *Structural equations with latent variables*. Wiley series in probability and mathematical statistics: Applied probability and statistics. Wiley-Interscience.

Bolton, L. and J. Alba (2006). Price fairness: Good and service differences and the role of vendor costs. *Journal of Consumer Research 33*(33), 258–265.

Bolton, L., L. Warlop, and J. Alba (2003). Consumer perceptions of price (un)fairness. *Journal of Consumer Research 29*(4), 474–491.

Bouch, A., A. Kuchinsky, and N. Bhatti (2000). Quality is in the eye of the beholder: Meeting users' requirements for Internet quality of service. *CHI Letters 2*(1), 297–304.

Bouch, A. and M. A. Sasse (1999). It ain't what you charge, it's the way that you do it: a user perspective of network QoS and pricing. In *Integrated Network Management VI. Distributed Management for the Networked Millennium. Proceedings of the Sixth IFIP/IEEE International Symposium on Integrated Network Management. (Cat. No.99EX302)*, pp. 639–654. IEEE.

Brennan, T. J. (2011). Net neutrality or minimum quality standards: Network effects vs . market power justifications. In I. Spiecker gen. Döhmann and J. Krämer (Eds.), *Network Neutrality and Open Access*, Baden Baden, pp. 1–21. Nomos Publishing.

Brown, J. D. (2001, April). Statistics corner. Questions and answers about language testing statistics: What is an eigenvalue? *Shiken: JALT Testing & Evaluation SIG Newsletter 5*(1), p.15–19. Retrieved from: http://jalt.org/test/bro_10.htm.

Browne, M. W. and R. Cudeck (1993). *Testing structural equation models*, Chapter Alternative Ways of Assessing Equation Model Fit, pp. 136–162. Bollen, K. and Long, J. S.

Byrne, B. (1989). *A primer of LISREL: Basic applications and programming for confirmatory factor analytic models.* Springer-Verlag Publishing.

Chau, C., Q. Wang, and D. Chiu (2010, March). On the viability of Paris metro pricing for communication and service networks. In *INFOCOM, 2010 Proceedings IEEE*, San Diego, California, USA, pp. 1–9. IEEE.

Cheng, H. K., S. Bandyopadhyay, and H. Guo (2011). The debate on net neutrality: A policy perspective. *Information Systems Research 22*, 60–82.

Chiou, J. (2004). The antecedents of consumers loyalty toward Internet Service Providers. *Information & Management 41*, 685–695.

Choi, J. P. and B. C. Kim (2010). Net neutrality and investment incentives. *RAND Journal of Economics 41*(3), 446–471.

Cisco (2012). Cisco visual networking index: Global mobile data traffic forecast update 2011. Technical report, Cisco. Retrieved from: http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.pdf.

Costello, A. and J. Osborne (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation 10*(7), 1–9.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika 16*(3), 297–334.

DeCoster, J. (1998). Overview of factor analysis. Accessed November 18, 2011, http://www.stat-help.com/notes.html.

Dischinger, M., M. Marcon, S. Guha, K. P. Gummadi, R. Mahajan, and S. Saroiu (2010). Glasnost: Enabling end users to detect traffic differentiation. In *Proceedings of the USENIX*

Symposium on Networked Systems Design and Implementation (NSDI), San Jose, CA, USA, pp. 1–14.

Dou, W. (2004). Will internet users pay for online content? *Journal of Advertising Research 44*(4), 349–359.

Economides, N. (2005). The economics of the Internet backbone. In S. K. Majumdar, I. Vogelsang, and M. E. Cave (Eds.), *Handbook of Telecommunications Economics*, Volume 2, pp. 373–412. Elsevier B.V.

Economides, N. and B. Hermalin (2012). The economics of network neutrality. *RAND Journal of Economics forthcoming*.

Economides, N. and J. Tag (2012). Net neutrality on the Internet: A two-sided market analysis. *Information Economics and Policy forthcoming*.

Farber, D. and M. Katz (2007, January). Hold off on net neutrality. Washington Post. Accessed April 26, 2010, http://www.washingtonpost.com/wp-dyn/content/article/2007/01/18/AR2007011801508.html.

Faulhaber, G. and D. Farber (2010). The open internet: A customer-centric framework. *International Journal of Communication 4*, 302–42.

Faulhaber, G. R. (2010). Transparency and broadband Internet service providers. *International Journal of Communication 4*, 738–757.

Faulhaber, G. R. (2011). Economics of net neutrality: A review. *Communications & Convergence Review 3*(1), 53–64.

FCC (2010, December). Report and order: In the matter of preserving the open Internet broadband industry practices. FCC 10-201.

Fehrenbacher, K. (2010, August 5). Caught on video: Google CEO dishes on Google Wave, Verizon & social strategy. Accessed March 29, 2011, http://gigaom.com/mobile/google-ceo-dishes-on-google-wave-verizon-social-strategy/.

Field, A. (2000). *Discovering statistics using SPSS for Windows: Advanced techniques for the beginner*. ISM (London, England). Sage Publications.

Fisher, R. (1993). Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research 20*(2), 303–315.

Fornell, C. and D. F. Larcker (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research 18*(1), 39–50.

Frey, B. S. and W. W. Pommerehne (1993). On the fairness of pricing - An empirical survey among the general population. *Journal of Economic Behavior & Organization 20*, 295–307.

Gaitonde, R. (2010, July). Chile legally mandates network neutrality and increased transparency for ISPs. Accessed August 05, 2010, http://broadbandbreakfast.com/2010/07/chile-legally-mandates-network-neutrality-and-increased-transparency-for-isps/.

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

Grimmelmann, J. (2010). Some skepticism about search neutrality. In B. Szoka and A. Marcus (Eds.), *The Next Digital Decade: Essays on the Future of the Internet*, Chapter 7, pp. 435–459. Washington, D.C.: TechFreedom.

Hagiu, A. (2006). Pricing and commitment by two-sided platforms. *The RAND Journal of Economics 37*(3), 720–737.

Hahn, R., R. Litan, and H. Singer (2007). The economics of wireless net neutrality. *Journal of Competition Law and Economics 3*, 399–451.

Hahn, R. and S. Wallsten (2006). The economics of net neutrality. *The Economists' Voice 3*(6), 1–7.

Hansell, S. (2008, August). Verizon's FiOS: A smart bet or a big mistake? New York Times. Accessed March 11, 2012, http://www.nytimes.com/2008/08/19/technology/19fios.html?pagewanted=2&_r=2&ref=technology.

Hermalin, B. and M. Katz (2007). The economics of product-line restrictions with an application to the network neutrality debate. *Information Economics and Policy 19*, 215–248.

Hinkin, T. R. (1998, January). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods 1*(1), 104–121.

Homans, G. (1961). *Social behavior: Its elementary forms*. Brace & World, New York, NY.

Hotelling, H. (1929). Stability in competition. *The Economic Journal 39*(153), 41–57.

Hoyle, R. (2011). *Structural equation modeling for social and personality psychology*. The Sage Library of Methods in Social and Personality Psychology. Sage Publications.

Hu, L. and P. Bentler (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal 6*(1), 1–55.

Jamison, M. and J. A. Hauge (2008). Getting what you pay for: Analyzing the net neutrality debate. mimeo, Social Science Research Network, Retrieved from: http://ssrn.com/abstract=1081690.

Janda, S., P. J. Trocchia, and K. P. Gwinner (2001). Consumer perceptions of Internet retail service quality. *International Journal of Service Industry Management 13*(5), 412–431.

Kilkki, K. (2008). Quality of experience in communications ecosystem. *Journal of Universal Computer Science 14*(5), 615–624.

Kim, S. and L. Stoel (2004). Apparel retailers: Website quality dimensions and satisfaction. *Journal of Retailing and Consumer Services 11*, 109–117.

Kleinrock, L. (1976). *Queuing systems, Volume 1.* New York: John Wiley & Sons, Inc.

Kline, R. B. (2005). *Principles and practice of structural equation modeling*. The Guilford Press.

Kocsis, V. and P. W. J. Bijl (2007). Network neutrality and the nature of competition between network operators. *International Economics and Economic Policy 4*(2), 159–184.

Krämer, J. and L. Wiewiorra (2012). Network neutrality and congestion sensitive content providers: Implications for content variety, broadband investment and regulation. *Information Systems Research forthcoming*.

Krämer, J., L. Wiewiorra, and C. Weinhardt (2012). Net neutrality: A progress report. *Submitted to Telecommunications Policy*, 1–48.

Lambert, P. (2010, February 24). Vodafone and Telefonica are overplaying their hand with Google. Accessed December 12, 2010, http://www.telecoms.com/18389/vodafone-and-telefonica-are-overplaying-their-hand-with-google/.

Lambrecht, A. and B. Skiera (2006a). Paying too much and being happy about it: Existence, causes and consequences of tariff-choice biases. *Journal of Marketing Research 43*(2), 212–223.

Lambrecht, A. and B. Skiera (2006b). Ursachen eines Flatrate-Bias: Systematisierung und Messung der Einflussfaktoren. *Zeitschrift für betriebswirtschaftliche Forschung 58*(8), 588–617.

Lee, R. and T. Wu (2009). Subsidizing creativity through network design: Zero-pricing and net neutrality. *The Journal of Economic Perspectives 23*(3), 61–76.

Lee, S. T. B. (2006, 05). Neutrality of the net. Accessed January 09, 2012, http://dig.csail.mit.edu/breadcrumbs/node/132.

Leiner, B. M., V. G. Cerf, D. D. Clark, R. E. Kahn, L. Kleinrock, D. C. Lynch, J. Postel, L. G. Roberts, and S. Wolff (2011). Brief history of the Internet. Accessed February 07, 2012, http://www.internetsociety.org/internet/internet-51/history-internet/brief-history-internet#REK78.

Lessig, L. (2001). *The future of ideas*. New York: Random House.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology 140*, 1–55.

MacCallum, R. and J. Austin (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology 51*(1), 201–226.

MacKie-Mason, J. and H. Varian (1995). Pricing congestible network resources. *IEEE Journal on Selected Areas in Communications 13*(7), 1141–1149.

Martín-Ruiz, D. and F. J. Rondán-Cataluña (2008). The nature and consequences of price unfairness in services: A comparison to tangible goods. *International Journal of Service Industry Management 19*(3), 325–352.

McDysan, D. (1999). *QoS and traffic management in IP and ATM networks*. McGraw-Hill, Inc. New York, NY, USA.

Meyer, D. (2011). Kroes attacks Dutch net-neutrality rules. Accessed February 11, 2012, http://www.zdnet.co.uk/news/regulation/2011/10/03/kroes-attacks-dutch-net-neutrality-rules-40094084/.

Mueller, M., A. Kuehn, S. M. Santoso, H. Asghari, B. Wagner, and X. Wang (2009). The network is aware - Social science research on deep packet inspection. The Network is Aware. Accessed January 14, 2012, http://dpi.ischool.syr.edu/MLab-Data.html.

Murray, A. S. (2010, April). Transcript: A conversation with Ivan Seidenberg. Council on Foreign Relations, Accessed March 24, 2012, http://www.cfr.org/technology-and-foreign-policy/conversation-ivan-seidenberg/p21840?breadcrumb=/publication/by_type/transcript.

Musacchio, J., G. Schwartz, and J. Walrand (2009, March). A two-sided market analysis of provider investment incentives with an application to the net-neutrality issue. *Review of Network Economics 8*(1), 22–39.

Nixon, P. (2011, March). Amended net net neutrality regulation calms critics. Business News Americas. Accessed March 24, 2012, http://www.bnamericas.com/news/telecommunications/amended-net-neutrality-regulation-calms-critics.

Njoroge, P., A. Ozdaglar, N. Stier-Moses, and G. Weintraub (2010). Investment in two sided markets and the net neutrality debate. Columbia Business School Working Paper, DRO-2010-05.

Nunnally, J. (1967). *Psychometric theory*. Tata McGraw-Hill Education, New York.

O'Brien, K. J. (2010, May). Pitting the web's users against its gatekeepers. New York Times. Accessed March 12, 2012, http://www.nytimes.com/2010/05/03/technology/internet/03neutral.html?pagewanted=1&src=busln.

O'Brien, K. J. (2011, June). Dutch lawmakers adopt net neutrality law. New York Times. Accessed August 21, 2011, http://www.nytimes.com/2011/06/23/technology/23neutral.html?pagewanted=all.

O'Connell, P. (2005, November 7). At SBC, It's all about "Scale and scope". Businessweek. Accessed July 23, 2010, http://www.businessweek.com/magazine/content/05_45/b3958092.htm.

Odlyzko, A. (1999, November). Paris metro pricing for the internet. In S. Feldman and M. Wellman (Eds.), *Proceedings of the 1st ACM conference on Electronic commerce*, Denver, CO, USA, pp. 140–147. ACM New York, NY, USA.

Odlyzko, A. (2009, March). Network neutrality, search neutrality, and the never-ending conflict between efficiency and fairness in markets. *Review of Network Economics 8*(1), 40–60.

Odlyzko, A., S. Hong, A. Pakanati, and V. K. Adhikari (2012). Minnesota Internet traffic studies (MINTS). New York Times. Accessed February 16, 2012, http://www.nytimes.com/2010/05/03/technology/internet/03neutral.html?pagewanted=1&src=busln.

Openmedia.ca (2011). Stop the meter. Openmedia. Accessed November 21, 2011, http://openmedia.ca/news/regulators-pull-back-usage-based-billing-after-half-million-canadians-speak-out.

Plusnet (2011). All about traffic management. Accessed January 19, 2011, http://www.plus.net/support/broadband/speed_guide/traffic_management.shtml.

Powell, S. (2006). Guru Interview: Sir Timothy Berners-Lee, KBE. Accessed November 11, 2011, http://first.emeraldinsight.com/interviews/pdf/berners-lee.pdf.

Rayburn, D. (2009). Deutsche Telekom enters the CDN market, partners with EdgeCast. Streamingmedia. Accessed March 19, 2012, http://blog.streamingmedia.com/the_business_ of_online_vi/2009/01/edgecast.html.

Reggiani, C. and T. Valletti (2012). Net neutrality and innovation at the core and at the edge. *The University of Manchester, The School of Economics Discussion Paper Series* (1202), 1–35.

Reitman, D. (1991). Endogenous quality differentiation in congested markets. *The Journal of Industrial Economics 39*(6), 621–647.

Renda, A. (2010, August). Neutrality and diversity in the Internet ecosystem. Centre for European Policy Studies (CEPS). Retrieved from: http://ssrn.com/abstract=1680446.

Rice, K. (2010, February). 60% market share and piercing vision make Akamai Technologies a tech stock survivor with massive potential, according to industry expert. Accessed February 02, 2012, http://www.twst.com/yagoo/KerryRiceInternetServicesTWO.html.

Rochet, J. and J. Tirole (2006). Two-sided markets: A progress report. *The RAND Journal of Economics 37*(3), 645–667.

Ronald, D. and R. L. Raschke (2007). An assessment of formative and reflective constructs in IS research. In *ECIS 2007 Proceedings*, pp. 1481–1492. University of St. Gallen.

Ronnen, U. (1991). Minimum quality standards, fixed costs, and competition. *The RAND Journal of Economics 22*(4), 490–504.

Sandvine (2010). Fall 2010 global Internet phenomena report. Technical report, Sandvine. Retrieved from: http://www.sandvine.com/downloads/documents/2010GlobalInternetPhenomenaReport.pdf.

Save the Internet (2011). What is net neutrality? Save the Internet. Accessed January 18, 2011, http://www.savetheinternet.com/faq.

Schneibel, G. and C. Farivar (2010, April 7). Deutsche Telekom moves against Apple, Google and net neutrality. Deutsche Welle. Accessed March 29, 2012, http://www.dw-world.de/dw/article/0,,5439525,00.html.

Schuett, F. (2010). Network neutrality: A survey of the economic literature. *Review of Network Economics 9*(1).

Schwartz, G., N. Shetty, and J. Walrand (2008). Impact of QoS on Internet user welfare. In C. Papadimitriou and S. Zhang (Eds.), *Internet and Network Economics*, Volume 5385 of *Lecture Notes in Computer Science*, Heidelberg, pp. 716–723. Springer.

Shaked, A. and J. Sutton (1982). Relaxing price competition through product differentiation. *The Review of Economic Studies 49*(1), 3–13.

Shin, D. (2007, July). User acceptance of mobile Internet: Implication for convergence technologies. *Interacting with Computers 19*(4), 472–483.

Sidak, J. G. and D. J. Teece (2010). Innovation spillovers and the "dirt road" fallacy: The intellectual bankruptcy of banning optional transactions for enhanced delivery over the Internet. *Journal of Competition Law and Economics 6*(3), 521–594.

Sluijs, J., F. Schuett, and B. Henze (2010). Transparency regulation as a remedy for network neutrality concerns: Experimental results. TILEC Discussion Paper No. 2010-039.

Statovci-Halimi, B. and G. Franzl (2011). QoS differentiation and Internet neutrality. *Telecommunication Systems*, 1–10.

Swanson, B. and G. Gilder (2008). Estimating the exaflood - The impact of video and rich media on the Internet. *Discovery Institute (Seattle, Washington).* Retrieved from: http://www.discovery.org/scripts/viewDB/filesDB-download.php?command=download&id=1475.

Sydell, L. (2007, September 19). Firms abandon online subscription plans. "All Things Considered" - National Public Radio, USA.

Teo, T., V. Lim, and R. Lai (1999). Intrinsic and extrinsic motivation in Internet usage. *Omega 27*(1), 25–37.

Vaidyanathan, R. and P. Aggarwal (2003). Who is the fairest of them all? An attributional approach to price fairness perceptions. *Journal of Business Research 56*(6), 453–463.

van Schewick, B. (2007). Towards an economic framework for network neutrality regulation. *Journal on Telecommunications & High Technology Law 5*(2), 329–392.

Varian, H. R. (2002, February). The demand for bandwidth: Evidence from the INDEX project. University of California, Berkeley. Mimeo. Retrieved from http://people.ischool.berkeley.edu/~hal/Papers/brookings.pdf.

Venkatesh, V., M. G. Morris, G. B. Davis, and F. D. Davis (2003). User acceptance of information technology: Toward a unified view. *Management Information Systems Quarterly 27*(3), 425–478.

Wagstaff, G. F. (1994). Equity, equality, and need : Three principles of justice or one ? An analysis of "Equity as desert". *Current Psychology 13*(2), 138–152.

Wallsten, S. (2007). Wireless net neutrality? Progress & Freedom Foundation Progress Snapshot Paper No. 3.2.

Walrand, J. (2008). Economic models of communication networks. In Z. Liu and C. H. Xia (Eds.), *Performance Modeling and Engineering*, Chapter 3, pp. 57–90. New York, USA: Springer.

Wang, Y., T. Tang, and J. Tang (2001). An instrument for measuring customer satisfaction toward web sites that market digital products and services. *Journal of Electronic Commerce*

Research 2(3), 89–102.

Wu, T. (2003). Network neutrality, broadband discrimination. *Journal on Telecommunications & High Technology Law 2*, 141–178.

Wu, T. (2007). Wireless Carterfone. *International Journal of Communication 1*, 389–426.

Wyatt, E. (2010, August 4). Google and verizon near deal on web pay tiers. New York Times, Accessed March 29, 2012, http://www.nytimes.com/2010/08/05/technology/05secret.html.

Xia, L., K. B. Monroe, and J. L. Cox (2004). This price is unfair! A conceptual framework of price fairness perceptions. *Journal of Marketing 68*(4), 1–15.

Yoo, C. (2005). Beyond network neutrality. *Harvard Journal of Law & Technology 19*, 1–77.

Zakon, R. H. (1997, November). Hobbes' internet timeline. Accessed January 16, 2012, http://tools.ietf.org/html/rfc2235.

Zinnbauer, M. and M. Eberl (2004). Die Überprüfung von Spezifikation und Güte von Strukturgleichungsmodellen: Verfahren und Anwendung. *Ludwig-Maximilians-Universität München: Schriften zur Empirischen Wirtschaftsforschung und Quantitativen Unternehmensplanung* (21), 1–27.