WP-11-01

ESMT White Paper

ASSESSMENT OF A SUSTAINABLE INTERNET MODEL FOR THE NEAR FUTURE

HANS W. FRIEDERISZICK, ESMT CA
JAKUB KAŁUŻNY, ESMT CA
SIMONE KOHNZ, ESMT CA
MICHAŁ GRAJEK, ESMT
LARS-HENDRIK RÖLLER, ESMT

ISSN 1866-4016

Citation

Friederiszick*+, H.W., J. Kałużny, S. Kohnz, M. Grajek, and L.-H. Röller (2011). *Assessment of a sustainable Internet model for the near future*. ESMT White Paper No. WP–11–01.

* Contact author: Hans W. Friederiszick, ESMT CA, Schlossplatz 1, 10178 Berlin,
Phone: +49 (0) 30 21231-7010, hans.friederiszick@esmt.org.

+ This report was commissioned by Deutsche Telekom. The opinions expressed herein are those of ESMT Competition Analysis (ESMT CA) and do not necessarily reflect the position of Deutsche Telekom. The conclusions are the results of the exercise of ESMT CA's best professional judgment.

Contents

1.	**Overview of results**	**6**
1.1.	Executive summary	6
1.2.	Extended summary	8
2.	**Introduction**	**19**
2.1.	Structure of the report	19
2.2.	Current business model of (eyeball) ISPs	20
2.3.	Stylized facts and resulting challenges	27
3.	**Typology of possible future business models**	**37**
3.1.	Dimensions of new business models	37
3.2.	Essential elements of future business models	43
3.3.	Conclusion on future business models	65
4.	**Theoretical analysis of effects**	**71**
4.1.	Economic first principles	72
4.2.	Detailed analysis of proposed business models	89
4.3.	Discussion of net neutrality definitions and their impact on possible market outcomes	118
Appendix		**124**
Economics literature on quality classes		124
Bibliography		128

Tables

Table 1: Relation between different business models and definitions of net neutrality 13

Table 2: Impact assessment of business models (relative to status quo business model) 16

Table 3: Overview over dimensions of business models 38

Table 4: Current business model (from an eyeball ISP's perspective) 42

Table 5: Differences between quality scenarios 46

Table 6: Pricing schemes for different commercial partners of eyeball ISPs 47

Table 7: Anchoring of “Congestion-Based Model” (BM 0) 66

Table 8: Anchoring of “Best Effort Plus” (BM 1) 67

Table 9: Anchoring of “Quality Classes – Content Pays” (BM 2) 69

Table 10: Anchoring of “Quality Classes – User Pays” (BM 3) 70

Table 11: Impact assessment of Congestion-Based Model (BM 0) 96

Table 12: Impact assessment of Best Effort Plus (BM 1) 106

Table 13: Impact assessment of Quality Classes – Content Pays (BM 2) 112

Table 14: Impact assessment of Quality Classes – User Pays (BM 3) 117

Table 15: Connection of BM with definitions of net neutralityutrality 120

Figures

Figure 1: Distribution of total Internet revenue, 2008 11

Figure 2: Simplified representation of the Internet architecture 25

Figure 3: Requirements of applications with respect to speed 30

Figure 4: Distribution of total revenue earned through non-business users by content providers, 2008 33

Figure 5: Distribution of total Internet revenue, 2008 34

Figure 6: Externalities between Internet participants 85

1. Overview of results

1.1. Executive summary

The increasing demand for bandwidth due to data-intense applications, the convergence of various digital communication technologies as well as the increasing commercial importance of the Internet has given rise to one of the most important questions in the coming years: whether and how the Internet economic model needs to evolve and what role regulation should play in this process. Net neutrality regulation, if and when formally implemented in some shape or form, has the potential to reallocate resources among industry participants, affect optimal pricing strategies and ultimately impact investment and innovation incentives. Through these effects, the regulatory framework is going to affect which business models will be at all feasible, which are going to thrive, and which will become obsolete.

The report derives and analyzes some likely future business models with a view to sustainability in terms of ability to accommodate increasing traffic volumes and social welfare implications. Based on these assessments the regulatory implications are discussed for each business model.

The stylized business models each focus on a different aspect: the “Congestion-Based Model” stresses the possibility to tackle congestion problems through congestion-based pricing. The “Best Effort Plus” preserves the traditional best effort network but gives ISPs more leeway with innovative services. The “Quality Classes – Content Pays” stresses the observed need of different applications for various degrees of quality of service. The “Quality Classes – User Pays” model, however, puts the focus on consumer choice for higher quality levels.

For the **“Congestion-Based Model,”** we find that it reduces congestion and allows more efficient utilization of the existing infrastructure. However, it is unlikely to provide sufficient incentives to entirely eliminate congestion. Still, it offers an increased participation of (light) users and more efficient incentives to invest in infrastructure due to better utilization. Content providers would be negatively affected in so far as they produce heavy traffic and cannot shape the traffic according to peak times. To the contrary, off-peak services (and investments in such services) could rise. From a broader policy perspective a minor drawback is that uncoordinated implementation can lead to increased complexity for content providers as well as subsequently end users.

In the **"Best Effort Plus"** scenario ISPs would offer premium services to content providers who need their content delivered at a premium rate (value added service). Guaranteed reserved bandwidth for priority novel services would ensure their quality or even viability, and thereby induce the creation of new services. Prices for best effort services are not expected to change. But in this scenario end users have additional access to separately marketed innovative services. However, the risk of foreclosure due to exclusive agreements and bundling strategies might be increased. This concern is alleviated within the European environment with its existing access regulation.

For the **"Quality Classes – Content Pays"** model, we find that higher qualities facilitate new content. Charging content rather than charging users for the higher quality levels is likely to maximize the value of the platform, and thereby increases incentives to invest both in infrastructure and content. In an environment with strong competition and the accordingly limited market power of individual ISPs in the best effort segment, the model creates little or no risk of foreclosure strategies. In the absence of effective competition, however, the model introduces a risk of under-investment into the infrastructure due to a strategic incentive: degrading quality in best effort might hike up the price for higher quality levels.

Like the business model discussed before, the **"Quality Classes – User Pays"** model also facilitates new content through higher qualities. However, charging users rather than the content provider for the higher quality levels is likely to lead to lower value and lower incentives to invest for the platform than in the previous business model. The regulatory risk related to foreclosure strategies seems smaller, though: the ability of a dominant ISP to favor a vertically-integrated content provider is lower. Both business models, "Quality Classes – Content Pays" and "Quality Classes – User Pays," bear a risk of market fragmentation in the event that no new Internet standards emerge.

The implementation of different forms of net neutrality regulation impacts the above business models to a different extent:

- The implementation of a **strong form of net neutrality** would prevent "Best Effort Plus" and "Quality Classes – Content Pays," but still allows the other two business models. This implies that some benefits of these new business models can be reaped with net neutrality regulation whereas other efficiencies cannot materialize: congestion-based pricing could decrease congestion to some extent and the ability to have differentiation quality classes for end users would open the possibility for higher quality content offerings. However, charging users rather than content providers for the higher quality levels is likely to lead to lower value and lower incentives to invest for the platform than a scenario where the content provider (also) pays. Furthermore, it might still be the case that delay-sensitive content is crowded-out of the network.

- In contrast, the implementation of a **weaker form of net neutrality** would enable the adoption of a business model which prices content providers for higher qualities. The comparison between content pays and user pays scenarios involves the following trade-off: the potentially increased risk of foreclosure in the content pays model must be weighed against inefficiency related to pricing the consumer side.
- Finally, the “Best Effort Plus” model implies that **net neutrality regulation only applies to traditional services** while novel innovative services would not be subjected to these rules. Ultimately, the crucial comparison is between this type of regulation versus a modest, but comprehensive net neutrality regulation. This comparison is, however, very complex and involves the quantification of effects as both models tend to increase the participation of end users and both open the way for a content-demanding higher quality of service.

As a consequence, in implementing the new EU regulatory framework for electronic communications, policy makers and regulators should carefully consider its impact on business models and the foregone benefits associated with those models in the short and long run. Since it is difficult to predict with any certainty which business models will dominate in the future, economic analysis suggests that authorities apply a patient “wait and see” approach: closely monitoring market developments and forcefully reacting to any emerging competitive threats rather than acting preemptively and therewith preventing some beneficial business models from developing.

1.2. Extended summary

The increasing demand for bandwidth due to data-intense applications, the convergence of various digital communication technologies as well as the increasing commercial importance of the Internet has given rise to one of the most important questions in the coming years: whether and how the Internet economic model needs to evolve and what role regulation should play in this process. An extensive debate in the US – including contributions by distinguished scholars – has been looking at the pros and cons of net neutrality regulation in the US context. Also in Europe, the European Commission’s consultation process in the second half of 2010, which resulted in over 300 responses, shows the vivid interest of policy makers and regulators, industry and the general public on that matter. However, what is missing is a thorough analysis of the implications of net neutrality regulation on some possible Internet business models adapted to the different market conditions in Europe, foremost European access regulation.

In this context, ESMT Competition Analysis analyzes the interaction between different net neutrality regulations and Internet business models. Net neutrality regulation, if and when formally implemented in some shape or form, has the potential to reallocate resources among industry participants, affect optimal pricing strategies and ultimately impact investment and innovation incentives. Through these effects, the regulatory framework is going to affect which business models will be at all feasible, which are going to thrive, and which will become obsolete. The report derives and analyzes some likely future business models with a view to sustainability in terms of the ability to accommodate increasing traffic volumes and social welfare implications. Based on these assessments, the regulatory implications are discussed for each business model.

The starting point is the exploration of eight fundamental features and developments of the industry features which will inevitably influence the future shape of the Internet.

Fact 1: Traffic is expected to increase significantly, in particular due to video-based applications. Actual traffic predictions predict that wired traffic will soar fourfold between 2009 and 2014. Video applications will contribute to a large extent to this growth as the share of Internet video alone will increase from about 30 percent of consumer Internet traffic to about 57 percent in 2014.

Fact 2: Over the course of the day, traffic volumes fluctuate greatly and high levels of congestion might be reached. In Europe, traffic peaks are being observed between 4:30 p.m. and 9:00 p.m. in 2010. In off-peak periods bandwidth utilization falls dramatically and lingers around one-fifth of peak capacity utilization between 1:00 a.m. and 5:00 a.m.

Fact 3: New applications such as 3DHD video, cloud gaming, and video conferencing require high-quality transmission standards. For example, streaming a YouTube video in HD quality requires 1.1 Mbit/s of transmission speed while streaming a 3D video in HD quality needs 50 Mbit/s. Increased needs for higher quality transmission are reflected by the increase in demand for quality of service enhancements provided by Content Delivery Networks such as Akamai and Limelight: revenues of CDNs specializing in video content are predicted to increase from below US$300 million in 2007 to over US$1.4 billion in 2012 representing an annual growth rate of 36 percent.

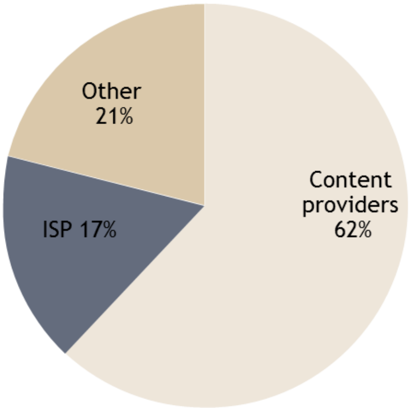
Fact 4: End consumers are currently priced so that they experience little or no incentive to control the traffic they generate. The OECD recent Global Communications Outlook states that “broadband also remains largely a flat-rate subscription in most countries.” Flat rates imply that the end users’ traffic consumption is largely unlimited and that heavy users are essentially subsidized by light users.

Fact 5: Peer-to-peer applications might jeopardize the payment balance under traditional transit agreements. P2P technology partially circumvents transit via lower-tier providers, thus reducing transit payments by content providers. At the same time, overall traffic is not reduced significantly. As P2P applications have gained importance in recent years, the amount content providers pay under transit agreements might no longer be a good approximation of the costs they produce on the entire network. To corroborate, in 2008 the peer-to-peer file sharing accounted for 32 percent of the total traffic on the Internet and for 22 percent of the global downstream traffic.

Fact 6: Network management practices allow a more cost-effective way to satisfy demand than over-provisioning. The increasing quality of transmission requirements of new applications like medical telemetry, network gaming, and video streaming require additional investments from the side of ISP. Whether the same quality of service has to be provided to all applications has a huge impact on the scope of investment: economic research finds that, to provide the same level of quality to new and traditional applications, ISPs would need to invest 60 percent more into infrastructure capacity than if differentiation in quality of service is allowed.

Fact 7: Content providers earn the largest share of the overall revenue in the Internet value chain. Content providers grab the largest share of the revenue earned on the Internet: in 2008, 62 percent of the total revenue was earned by content and service providers, while Internet service providers cashed only 17 percent.

Figure 1: Distribution of total Internet revenue, 2008



Source: AT Kearney (2010).

Fact 8: The segment of content providers becomes increasingly concentrated. The Internet becomes a more and more concentrated economic system with a relatively small number of participants (hosting, cloud and content providers) accounting for the increasing share of the total traffic: "out of the 40,000 routed end sites in the Internet, 30 large companies – 'hyper giants' like Limelight, Facebook, Google, Microsoft and YouTube – now generate and consume a disproportionate 30 percent of all Internet traffic." With an increasingly concentrated content provider side, it can be expected that the share of the jointly generated surplus that ISPs can appropriate is going to deteriorate.

These developments of the Internet indicate that the current business model might not be sustainable in the future and that changing to more tailored business models might open new opportunities for ISPs. On the basis of those facts, the report derives alternative Internet business models from the point of view of ISPs' profit maximization. Each one of the business models focuses on a different aspect.

The first business model, named "**Congestion-Based Mode**l," stresses the possibility to tackle congestion problems through congestion-based pricing, however no quality differentiation is introduced. Specifically, in this business model ISPs are assumed to charge content providers higher prices for traffic in peak periods than in off-peak periods. For example, the cost for a provider of movie downloads of an end user downloading an HD movie during the peak evening period could be significantly higher than if the same movie was downloaded in the early morning hours or within a 24-hour period. End users in this business model can choose between flat rates with differentiated data caps.

The second model, named **"Best Effort Plus,"** considers a two-tiered Internet structure. It preserves the traditional best effort network for traditional (existing) services and assumes that content providers and end users are priced as in the status quo if they operate on the best effort level. However, these restrictions do not apply to innovative future services, for which pricing and guaranteed service requirements follow individual negotiations between the eyeball ISP and the content provider. So the Internet as we know it would keep operating under similar principles as it does today but there would be more flexibility in the provision of novel services and the pricing thereof. For example, an ISP could charge a premium price from an innovative e-health service provider in return for guaranteeing a specified level of transmission quality (premium service). This model implies that a greater level of vertical cooperation between ISPs and content providers is necessary to implement quality guarantees. Future innovative services would remain unregulated; however, policy makers and regulators would have to define what defines an innovative service and which type of service is thus exempted from net neutrality regulation.

The third model, labeled "**Quality Classes** – **Content Pays,**" stresses the perceived need of different applications for various degrees of quality of service and offers different quality classes open for different applications. Unlike in the previous business model, the quality classes encompass all services, including currently available traditional services. Depending on their requirements, content providers could purchase the transit quality most appropriate for its type of content. For example, a content provider offering HD movie streaming or gaming services requiring low latency would purchase a more expensive premium quality class to ensure the quality of experience for end users. In contrast, for delivering an e-mail a cheaper, lower priority class could be chosen. It would become the ISPs responsibility to deliver the quality of service paid for by the content provider. In other words, content providers could choose to pay a premium price for a higher quality of transmission of their data. End users would still pay a uniform flat rate in this model and experience the quality as chosen by the content provider.

The last model, labeled "**Quality Classes** – **User Pays,**" however, puts the focus on consumer choice for higher quality levels and offers multiple quality classes for end users that are designed to match their different usage patterns. For example, end users who frequently use interactive applications might choose the quality class which is more suitable for dealing with such applications, i.e., that offers a low level of delay and jitter. Other users, who focus on multimedia applications, might choose another quality offering characterized by low packet loss and high bandwidth, and so on.

Some of the business models considered are more "net neutral" than others and the following table illustrates the relationship between the different business models and definitions of net neutrality, which have been put forward by proponents of net neutrality in the public discussion, in particular in the US. They serve as a reference for the academic evaluation of the competitive effects for the stylized business models.

Table 1: Relation between different business models and definitions of net neutrality

	Def 1: Equal treatment	Def 2: Reasonable network management, but no payment for quality by CP (status quo)	Def 3: Non-discriminatory quality classes	Def 4: No termination fees
BM 0: Congestion-Based Model	Violated*	Not violated	Not violated	Not violated
BM 1: Best Effort Plus	Violated	Violated	Violated	Violated
BM 2: Quality Classes – Content Pays	Violated	Violated	Not violated	Not violated
BM 3: Quality Classes – User Pays	Violated	Not violated	Not violated	Not violated

Note: *Also the “Congestion-Based Model” assumes that reasonable network management is undertaken
Source: ESMT CA.

In the following section, we then summarize and characterize seven general and robust results from the economic literature that are relevant for the assessment of the expected effects of new business models from the **social welfare point of view**.

Principle 1: Common-pool resources are characterized by congestion and suboptimal levels of investment. A number of fundamental design features allow treating the Internet as a common-pool resource and hence make the existence of problems typical for common-pool resources likely. Common-pool resources are characterized by difficulties developing physical or institutional means of excluding beneficiaries (so-called non-excludability). This leads to strong temptations to free ride and consequently to suboptimal investment in the resource. At the same time, common-pool resources share with private goods that one person’s consumption subtracts from the quantity available to others (so-called rivalry). This implies that common-pool resources experience congestion problems unless use limits are devised and enforced.

Principle 2: Product differentiation increases total welfare. The introduction of product differentiation quite generally generates positive welfare effects. Broadly speaking, product differentiation increases welfare because it increases the number of available choices and allows heterogeneous consumers to choose consumption bundles more closely suited to their individual preferences.

Principle 3: Price discrimination increases total welfare. Price discrimination describes a practice of charging different buyers different net prices for the same product. Although price discrimination may invoke negative reactions and connotations among the public, it is a practice that is widespread in a variety of market settings. A common understanding among the economic profession is that it is generally welfare-enhancing and price discrimination only occasionally raises competition concerns.

Principle 4: A price increase to content providers reduces the price to end users ("waterbed effect"). A well-established and quite general theoretical result in the literature on two-sided markets states that increasing prices for one side usually leads to lower prices for the other side. This effect has important implications for net neutrality regulation: in such a setting, the allocative effect of higher charges on content providers implies a (partial) transfer from content providers toward end users.

Principle 5: The difference in expected profitability with and without investment/innovation affects incentives to invest and innovate. Expected profitability depends to a large extent on the competitive environment. Uncontested monopolists have low incentives to invest and innovate in their core markets (so-called *fat-cat effect*). If industry participants expect competitive conditions in the future, they will also have low incentives to invest and innovate because they expect that profits from their innovation are going to be competed away. Incentives to innovate are largest in highly contestable or oligopoly markets. Innovations allow firms to differentiate from each other and thus lessen competitive pressure or prevent rivals from "catching up." Strategic considerations may provide additional incentives to invest, for example, to deter entry or the expansion of rivals.

Principle 6: Network industries benefit from interoperability. Network effects are similar to economies of scale: as the number of buyers and sellers both increase, the surplus available to each agent also increases. Therefore, the more members a network attracts the more value it generates for its members. Also, network effects often involve externalities in the sense that prices do not fully incorporate the benefits of one person's entry into the network on existing members. This leads to the under-adoption of the network. Interoperability between different networks increases the size of the overall network available to end users and hence increases welfare.

Principle 7: Economic decisions involve trade-offs. Economic decisions usually involve making a trade-off. This also applies to regulatory decisions which affect how business is carried out on the Internet. Some important trade-offs are discussed in the report:

- Consumer benefits from lower prices today versus consumer benefits from new content-related products and services tomorrow
- High quality of service for some versus average quality of service for all content providers or end users
- Incentives to innovate in content and services vs. incentives to invest/innovate in infrastructure provision (for non-complementary network and content investments)
- Net benefits of ex ante versus ex post regulation (antitrust enforcement)

Based on these economic first principles and trade-offs we identify the major social benefits and costs linked to each business model and discuss regulatory options in relation to the different business models. Each of the business models may lead to a different overall welfare implication (e.g., increasing overall efficiency or utilization of the infrastructure) as well as to different financial transfers across market participants (e.g., from content providers to ISPs, or vice versa). To the extent that there are asymmetries in the geographic distribution of different players (e.g., many large content providers are located in the US), financial transfers across market participants may also imply financial transfers across different world regions (e.g., from Europe to the US, or vice versa). The main social benefits and costs linked to each business model are summarized in the following table.

Table 2: Impact assessment of business models (relative to status quo business model)

		BM 0: Congestion Based	BM 1: Best Effort Plus	BM 2: Quality Classes – Content Pays	BM 3: Quality Classes – User Pays
Congestion		Likely decrease	Persistent congestion on best effort with limited spillovers from innovative services	Reduction for high quality, but increase for the low quality (in the short run)	Reduction for high quality, but increase for the low quality (in the short run)
End users	Prices	Heavy users may pay more, but light users less	Unchanged in best effort, independent marketing of innovative services	Possibly reduction due to the waterbed effect	Availability of inexpensive low-quality flat rates, but increased flat rate for premium class
End users	Participation	Increase	No change for best effort, participation in innovative services	Increase	Increase due to e.g., inexpensive low-quality flat rates, but decrease due to waterbed effect
Content provider	Prices	Increase for CP unable to shape traffic and decrease for other CPs	Unchanged for best effort and increase for premium class	Unchanged for best effort and increase for premium class	Possibly reduction due to the waterbed effect
Content provider	Participation	Less crowding-out of delay-sensitive services, but some exit of CPs unable to steer traffic	Limited effects for best effort class and premium services are facilitated	Increase due to facilitation of premium services, but decrease due to waterbed effect	Increase due to facilitation of premium services and waterbed effect
ISPs' investment		More efficient investment incentives	Increased incentives to invest in capacity for innovative services, no (limited) incentives to under-invest in best effort	Increased incentives to invest due to quality differentiation, limited incentive to degrade best effort quality	Increased incentives to invest due to quality differentiation, limited incentive to degrade best effort quality
Regulatory costs	Ex ante	Transparency requirements	Transparency requirements Regulation and monitoring of innovative services	Transparency requirements Potential introduction of minimum quality of service	Transparency requirements
Regulatory costs	Ex post	No increase in the risk of foreclosure	Increase in the risk of foreclosure (discriminatory access to high quality class, exclusivity, bundling)	Increase in the risk of foreclosure (discriminatory access to high quality classes)	No increase in the risk of foreclosure
Interoperability		Limited coordination problems	Fragmentation risk regarding innovative services	Fragmentation risk concerning higher quality classes	Fragmentation risk concerning higher quality classes

Source: ESMT CA.

For the **"Congestion-Based Model,"** we find that it reduces congestion and allows a more efficient utilization of the existing infrastructure. However, it is unlikely to provide sufficient incentives to entirely eliminate congestion. Still, it offers an increased participation of (light) users and increased incentives to invest in infrastructure due to better utilization (which does not necessarily result in more investment relative to a counterfactual without peak-load pricing as peak traffic demands are smoothed). Content providers will be negatively affected in so far as they produce heavy traffic and cannot shape the traffic according to peak times. To the contrary, off-peak services (and investments in such services) could rise. From a broader policy perspective a minor drawback is that uncoordinated implementation can lead to increased complexity for content providers as well as subsequently end users.

In the **"Best Effort Plus"** scenario ISPs gain the option to offer premium services to content providers who need their content delivered at a premium rate (value added service). Guaranteed reserved bandwidth for priority novel services would ensure their quality or even viability, and thereby induce the creation of new services. Prices for best effort services are not expected to change. However, end users have additional access to separately marketed innovative services. However, the risk of foreclosure due to exclusive agreements and bundling strategies might be increased. This concern is alleviated within the European environment with its existing access regulation.

For the **"Quality Classes – Content Pays"** model, we find that higher qualities facilitate new content. Charging content rather than users for the higher quality levels is likely to maximize the value of the platform, and thereby increases incentives to invest both in infrastructure and content. The model, however, introduces a risk of under-investment into the infrastructure due to a strategic incentive: degrading quality in best effort might hike up the price for higher quality levels. The effect is substantially reduced or even eliminated, though, in an environment with limited market power of individual ISPs in the best effort segment. In so far as the model proves to be problematic, a minimum quality of standard regulation might be required.

Like the business model previously discussed, the **"Quality Classes – User Pays"** model also facilitates new content through higher qualities. However, charging users rather than the content provider for the higher quality levels is likely to lead to lower value and lower incentives to invest for the platform than in the previous business model. The regulatory risk related to foreclosure strategies seems smaller though: the ability of a dominant ISP to favor a vertically-integrated content provider is lower. Both business models, "Quality Classes - Content Pays" and

"Quality Classes – User Pays," bear the risk of fragmentation in so far as no common Internet standard emerges.

The implementation of different forms of net neutrality regulation impact the above business models to a different extent:

- The implementation of a **strong form of net neutrality** prevents "Best Effort Plus" and "Quality Classes – Content Pays," but still allows the other two business models. This implies that some benefits of these new business models can be reaped with net neutrality regulation whereas other efficiencies cannot materialize: congestion-based pricing could decrease congestion to some extent and the ability to have differentiation quality classes for end users would open the possibility for higher quality content offerings. However, charging users rather than content providers for the higher quality levels is likely to lead to lower value and lower incentives to invest for the platform than a scenario where the content provider (also) pays. Furthermore, it might still be the case that delay-sensitive content is crowded-out of the network.
- In contrast, the implementation of a **weaker form of net neutrality** would enable the adoption of a business model which prices content providers for higher qualities. The comparison between content pays and user pays scenarios involves the following trade-off: the increased risk of foreclosure in the content pays model must be weighed against inefficiency related to pricing the consumer side.
- Finally, under the "Best Effort Plus" model any **net neutrality regulation could only apply to traditional services** while novel innovative services would not be subjected to these rules. Ultimately, the crucial comparison is between this type of regulation versus a modest, but comprehensive net neutrality regulation. This comparison is, however, very complex and involves the quantification of effects as both models tend to increase participation of end users and both open the way for content demanding higher quality of service.

As a consequence, in implementing the new EU regulatory framework for electronic communications, policy makers and regulators should carefully consider its impact on business models and the foregone benefits associated with those models in the short and long run. Since it is difficult to predict with any certainty which business models will dominate in the future, economic analysis suggests that authorities apply a patient "wait and see" approach: closely monitoring market developments and forcefully reacting to any emerging competitive threats rather than acting preemptively and therewith preventing some beneficial business models from developing.

2. Introduction

The increasing demand for bandwidth due to data-intense applications, the convergence of various digital communication technologies as well as the increasing commercial importance of the Internet has given rise to one of the most important questions in the coming years: whether and how the Internet needs to be regulated. Within that context ESMT Competition Analysis has been asked to analyze various business models including the current one with a view to sustainability in terms of the ability to accommodate future traffic flows and social welfare implications. Furthermore, the report analyzes the potential regulatory implications resulting from different prevailing business models.

2.1. Structure of the report

The report addresses the questions on the sustainability and regulatory impact of different business models by first illustrating the current prevailing business model vis-à-vis all relevant commercial partners; that is, content providers, interconnecting ISPs and end users. It then proceeds to highlight some fundamental facts which result from this business model in conjunction with the changing nature of the Internet. These fundamental facts provide some background as to why the discussion of alternative business models is interesting: when discussing future business models, one should keep in mind the question of why ISPs have not established the proposed new business models already. To answer this, it is important to realize the fundamental changes which have occurred in the past.

Section **3** then identifies alternative Internet business models. First, we span the different dimensions of a complete business model. A business model covers the commercial relation between the ISP and all relevant players in the market. It also includes the level of vertical integration that the ISP is envisaging as this in turn determines the type of commercial relationship between different partners in terms of supplier or competitor. The resulting space for alternative business models is large. Essentially, each dimension can be crossed with all other dimensions in order to generate potential business models. In a second step, we thus identify elements of viable business models by analyzing them from the point of view of ISPs' profit maximization. Finally, we summarize the arguments in relation to elements of viable business models by proposing four different complete business models. The last part thus offers a description of a number of different business models which could emerge or could become more prominent in the future. Each one of the business models focuses on a different aspect: the first business model stresses the possibility to tackle congestion problems through congestion-based pricing. The second model preserves the traditional best effort network but gives ISPs more leeway with innovative services. The third model stresses the perceived need of different applications for various degrees of quality of service. The last model, however, puts the focus on consumer choice for higher quality levels.

The next section analyzes the welfare effects of the proposed business models. In comparison to the previous section, which focuses on the profit prospects, it thus assumes the perspective of a social planner who evaluates different future scenarios. The first part establishes economic first principles which can be used in order to evaluate the different business models. Those are general overarching principles in economics which help us align the detailed discussion in the second part of the section. Since the business models would be subject to the future net neutrality regulation which is currently being debated on the EU level, the discussion of the business models has to be seen within this context. Net neutrality regulation, if and when formally implemented in some shape or form, has the potential to reallocate resources among industry participants, affect optimal pricing strategies and affect their investment and innovation incentives. Through these effects, the shape of net neutrality regulation is going to affect which business models are going to be at all feasible, which are going to thrive, and which are going to become obsolete. Therefore, when assessing the potential impact of net neutrality regulation, one needs to consider how the regulation may affect future business models. The last section thus discusses regulatory options in relation to the different business models.

2.2. Current business model of (eyeball) ISPs

Currently, ISPs maintain direct contractual relations to end customers (also referred to as end users or subscribers), directly connected content providers and other Internet access and service providers (ISPs) providing interconnection services.[1] Eyeball ISPs predominantly provide access to end users (eyeballs) whereas content ISPs focus on the provision of access to content providers. In the following, we provide some background of the relations between those agents from the point of view of an eyeball ISP.

---

1 In the following, we simplify the exposition by using the term ISP for providers of Internet infrastructure on all levels. It thus subsumes providers of all tiers including last-mile providers.

2.2.1. End users

End customers use the Internet for various activities like social networking, games, e-mail, instant messaging, multimedia, research, or e-commerce.[2] Currently, end users are typically charged flat rates which might be differentiated according to maximum bandwidth of the offered connection. Also, in some OECD countries flat rates are complemented with data caps. Bundling of flat rates for Internet, TV, and fixed mobile telephony is also an aspect of end user pricing.

2.2.2. Content providers

Commercial content providers offer content or services on the Internet.[3] This does not necessarily mean that content providers themselves own the rights for the offered content. In fact, content ownership rights are often not carried by the content provider. In order to avoid confusion, we will refer to content providers as the companies offering content and services on the Internet and to content rights owners as those companies which hold the rights over the published content.

Content providers earn their revenues predominantly by charging sellers of products and services for online advertisement, by online transactions (e-commerce) and/or subscription fees on online services. Some content providers earn their revenues mainly by charging for complementary offline products or services (e.g., newspapers, educational institutions, and many others). See section **2.3** for detailed information on the distribution of revenue types for content providers.

Currently, commercial content providers pay the ISP which is directly connected to them for placing their content on the Internet and making it accessible to the end user (access charge). Pricing schemes for access only to small content providers are typically flat rates differentiated to first-mile bandwidth.[4] In relation to mid-sized and large content providers, there exists an array of different contractual

---

2 See for example the Nielsen reports (http://blog.nielsen.com/nielsenwire/online_mobile/what-americans-do-online-social-media-and-games-dominate-activity/).

3 It should be noted that the term end customer or subscriber is not defined unambiguously as end customers might also provide content on the net e.g., in the form of personal webpages. Therefore, it is important to distinguish content providers who provide content with a commercial perspective. Commercial content comprises not only offers with a user-oriented network interface ("web pages"), but also databases, applications or complete systems with a machine-oriented Internet interface (see for example the offerings under http://aws.amazon.com/).

4 Often flat-rate agreements are combined with hosting contracts. See for example Host Europe, one of the major European webhosting providers.

solutions. Typically, those content providers are priced according to some volume-related pricing scheme. In particular, interconnection agreements between content providers and ISPs often specify payments according to the 95th percentile traffic volume that they generate over a 30-day period. The 95th percentile measurement samples traffic volumes on a five-minute basis over the entire month. At the end of the month, the five-minute volume samples are ordered lowest to highest and payment is according to the product of the 95th percentile of the volume samples (Mbps) and the transit price (EUR/Mbps). The underlying idea is that each content provider is charged for the entire month according to the maximum bandwidth that it used during the month. However, the maximum is measured at the 95th percentile assuming that the remaining 5th percentile highest volume samples represent extraordinary traffic which is not included in the pricing. However, some (very large) content providers which are vertically integrated into infrastructure might even enjoy peering agreements. Content providers do not directly pay any charges to ISPs which are not directly connected to them.[5]

2.2.3. Other ISPs

ISPs provide the infrastructure on various layers of the Internet. Originally, the Internet was organized in a tripartite hierarchy with local networks (tier 3), regional ISP networks (tier 2) and the Internet backbone (tier 1).[6] A typical exchange between an end user and a content provider within this tripartite structure required uploads and downloads across various layers of the network. For example, a content provider contracted with a particular first-mile (local) provider. This provider delivered the content up to the next connection node of the regional ISP which in turn passed the traffic to the backbone provider who delivered it to the regional ISP and then to the last-mile provider of the end user. Accordingly, in order to be able to provide end-to-end services to customers, the different layers had to connect in order to route the information packets. The interconnectivity was provided by negotiating peering or transit agreements.

In *peering agreements*, two ISPs agree to exchange traffic between each other and their respective customers without money transfers. Peering agreements do not

---

5 It should be noted that larger content providers might be vertically integrated into the infrastructure (forming so-called autonomous systems, e.g., Google, Facebook, Amazon or eBay). Such content providers might enter into peering and transit agreements with other autonomous systems or ISPs. This makes a strict separation between a content provider and an infrastructure provider difficult – a fact that should be kept in mind for the entire assessment.

6 A typical definition of a tier 1 provider focuses on the fact that the operator can provide access to the entire Internet by only peering with other operators.

include transit services to the networks of a third party. Peering agreements typically specify a ratio of inbound and outbound traffic in order to ensure that the contract is economical for both sides. In the early days of the privatized Internet only tier 1 providers engaged in peering agreements. Whether an ISP decides to engage in peering or transit depends on a number of criteria, such as geographical coverage, customer size and mix, network quality, etc. Peering requires a direct connection which comes at a cost. These costs have to be outweighed by the benefits of peering (i.e., reduced transit fees and reduction in the number of hops). Typically, comparably-sized systems decide to peer with each other whereas more asymmetric systems implement transit agreements. This implies that tier 1 operators typically peer with each other but have transit agreements with tier 2 operators, and vice versa.[7]

In contrast, *transit agreements* imply that one ISP pays another ISP for delivering its traffic. Transit agreements typically specify a price for peak bandwidth utilization (megabit per second per month – often 95 percent of measured peak bandwidth utilization is billed).[8] Transit agreements typically offer delivery within the entire Internet; this means, the receiving ISP has to assure delivery over third-party networks.[9]

This simple tripartite structure has changed significantly within the last decade due to several developments such as multi-homing or secondary peering.[10] These developments influenced, on the one hand, the hierarchical structure of the Internet, and on the other hand the pricing schemes and money flows between the different players. Furthermore, they have implications for net neutrality: the move from a strict hierarchical system to a more complex interconnection map, for example, through secondary peering, can imply that essentially identical traffic is delivered with different speeds and levels of quality, and for those services different prices can be paid.

With multi-homing tier 2, providers started to directly connect to more than one tier 1 provider in order to secure connection in case of the failure of one of those upstream providers as well as mitigating the market power of single providers.

---

7 Two ISPs can also decide to de-peer if one party realizes that it does not benefit economically from the peering agreement.

8 See the previous section for more details on the 95th percentile measurement.

9 Usually, peering and transit agreements also include the so-called Service Level Agreements (SLA) which specify indicators related to the quality of delivery over one's own network or certain regions. The most important SLA indicators are latency (speed) and packet loss.

10 Much of the following exposition relies on the recent paper by Yoo (2010).

Furthermore, beyond strategic motivations multi-homing is also employed by ISPs on various levels to improve general transmission efficiency.

Furthermore, peering used to take place mainly between the early backbone providers. However, tier 2 providers started to connect directly to each other in order to avoid transit fees of upstream providers and increase the quality of the connection (this is known as secondary peering). The need for partial transit arose due to the fact that second-tier providers started to peer with each other. They, thus, did not need to buy transit for the entire Internet, but only for parts of it (those parts with which they were not already peering).[11]

When two ISPs find that on balance one of the two peers derives a higher benefit from the peering agreement, they might decide that one side will have to pay for the connection of its customers, but not the other way around.[12] Over the years, two types of different last-mile providers have emerged, those that provide access to mainly content providers (so-called “content networks”) and those that provide access mainly to end users (so-called “eyeball networks”). Since the costs of eyeball networks tend to be significantly higher than the costs of content networks, paid peering occurs so that content networks pay for the peering with eyeball networks.

Applications started to be based on a peer-to-peer instead of a client-server architecture. This implies that devices at the edge of the network are no longer divided between devices that primarily host content and those that primarily request content. Instead, both devices perform both functions. This development has implications for the capacity at different layers of the Internet, for example, more uploading bandwidth is necessary for the receiving and hosting client at the last mile. Furthermore, this development also impacts on the traditional payment scheme where end users paid a flat rate and content providers paid for maximum volumes.[13]

---

11 Additionally, partial transit agreements can take the form of selling only outbound traffic in case of traffic ratios which exceed the predefined levels within the peering agreements.

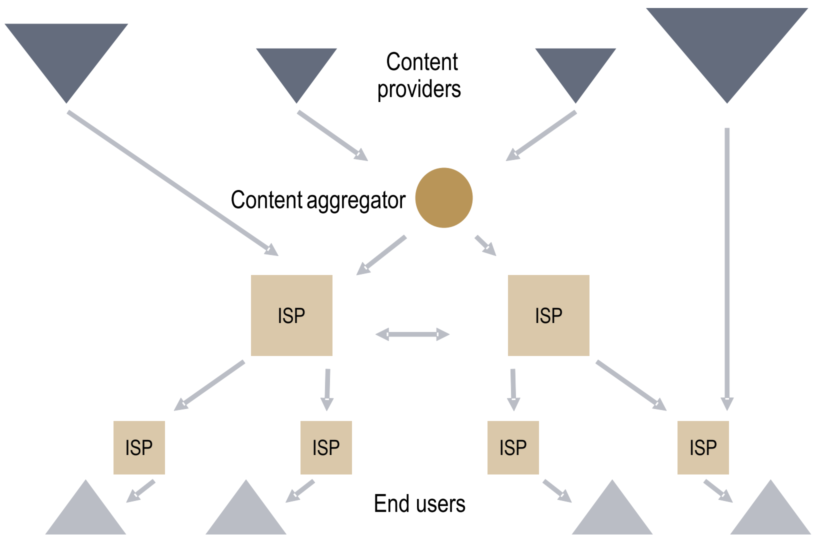
12 The difference with transit agreements is that the system that sells its connectivity does not provide transit to the rest of the Internet.

13 Take the example of 10 end users wanting to access a video of 100 megabytes. With the traditional client-server structure, each end user would have to request the 100 megabytes from the server of the content provider. In this stylized example this would result in a necessity of maximum volume of 1,000 megabyte for the content provider. With the change to a P2P structure, it can be sufficient that one of the end users downloads the movie and the rest of the end users receive the movie from their peer. This implies that the maximum volume for the content provider can be reduced to 100

An assessment of the sustainability of the current pricing system has to take those structural changes of the Internet hierarchy into account. In particular, it should entail an assessment of how traffic is affected by those changes in the future and what this implies for the different payment streams for different players.

Figure 2 illustrates the basic architecture of the Internet depicting the end users, local and transit ISPs and content providers. For the purpose of simplicity and in order to illustrate the two-sidedness of the market, the illustration collapses the notion of pure backbone providers and regional providers into transit ISPs that tend to connect content producers whereas local ISPs connect the end users. The reason for differentiation between those types of ISPs is that, typically, local ISPs due to the high fixed costs – and the resulting substantial economics of scale – of connecting individual end users are thought of as having some market power. In comparison, backbone providers or ISPs connecting content providers are typically thought of as having little to no market power. The idea is that commercial content providers are more flexible when selecting their respective connection to the Internet as they are – in terms of the location – less bound to the local provider. Furthermore, scale economies play a lesser role when connecting a content provider which generates large amounts of traffic individually.

Figure 2: Simplified representation of the Internet architecture



Source: ESMT CA.

---

megabytes and, thus, its transit payments to the upstream ISP can be significantly reduced. As consumers have flat rates, the network operators will receive lower payments for similar data transfers (absent a change in the fixed fee).

The above illustration of the Internet architecture highlights that different content providers can be connected in different ways to the local ISP and therewith to the end users: they can either be directly connected to the local ISP or indirectly over a transit ISP or a content aggregator. The depiction therefore also takes into account the following changes in the Internet structure:

- **Server farms providing a direct link to local ISPs**: In order to reduce costs and improve latency, large content providers have started to deploy server farms instead of single large servers. This reduces the transit paid to ISPs as less traffic has to be routed through the backbone itself. In a recent study, Gill et al (2008) estimate that for the top three content providers (Google, Yahoo and Microsoft) about one-third of their traffic is routed without using the Internet backbone.
- **Content delivery networks (CDN) serving as content aggregators:** Content providers might pay for quality of service and higher reliability by using CDNs. CDNs substitute transit traffic by storing content in a network of local caches. This enables CDNs to serve several requests for the same content from the one cache, thereby saving on network resources.[14]

In terms of the current commercial interaction, the situation can be summarized as follows: (1) End users pay a flat rate differentiated with respect to certain bandwidth for access to the Internet to their connecting ISP. (2) Small content providers pay flat access rates whereas larger content providers pay according to maximum volumes generated over the month. Large content providers who are integrated into the infrastructure might have peering agreements with local ISPs. (3) Content aggregators and other interconnecting ISPs typically have either peering or transit agreements.

2.2.4. Quality of transmission and traffic management

The uniform quality of transmission in the current Internet is essentially based on best effort, which is applied to all traffic. Best effort incorporates the idea that

---

[14] A CDN requests data from a contracted content provider once and stores it in several local caches in order to deliver it to multiple local end users. The following stylized example illustrates how this process decreases (long distance) transit payments for the content provider. Suppose 10 end users in country A want to view content from a content provider in country B. Each direct data request from the end user to the content provider would imply that the data has to transit from country B to country A, incurring transit costs. If, however, the end user is redirected to the local cache in country A, then for the individual 10 data requests only one long distance transit occurs.

the given infrastructure (at a given moment in time) is allocated to the entire traffic in an (essentially) egalitarian way.

Traffic management tools may be employed to prioritize more time-sensitive applications such as IPTV and streaming video. Network operators already implement network management techniques that allow the prioritization of traffic based on (the delay sensitivity of) the application type. For example, standard internet routers already prioritize traffic according to four quality classes in order to smooth temporary bottlenecks that result within intervals of milli- or microseconds.[15] Those management practices are designed to avoid isolated events of very short-termed strains (so-called “micro bursts”) on the Internet capacity. According to Deutsche Telekom, this form of (micro-) prioritization is not sufficient to keep the system stable and packet loss at a minimum if traffic is consistently close or over-capacity.[16]

Furthermore, in some instances ISPs have also throttled the traffic of certain heavy users – often this relates to users heavily engaging in P2P activities – in order to stabilize the overall system and/or to enforce the contractual conditions of usage (“fair use” policies).[17] But so far operators, having no pecuniary incentives for doing more, have been using these techniques to a very limited extent in order to ensure a minimum of connectivity to all users, but not to achieve full efficiency of their existent network resources according to their usage.

2.3. Stylized facts and resulting challenges

In our analysis we take into account main stylized facts which describe crucial developments and features of the industry. These developments and features will inevitably influence the future shape of the industry as well as the outcome of any policy. Below, we list the facts and support each of them with corresponding empirical evidence.

Fact 1: Traffic is expected to increase significantly, in particular due to video-based applications.

---

15 The four quality classes are attributed according to certain characteristics of the particular traffic. Essentially, delay-sensitive traffic (such as VoIP) is differentiated from delay-insensitive traffic (such as e-mail).

16 In fact, it appears to be unclear how the overall system would react to a significant period of over-utilization of capacity.

17 For a detailed discussion of Internet congestion and countermeasures, see Bauer et al (2009).

Fundamental drivers of the Internet usage indicate the steady growth of the traffic in the past and give an idea of the projected development in the future. The number of Internet users rose from a paltry 16 million in 1995 to 1.7 billion in 2009 representing an annual growth rate of 40 percent.[18] This high growth rate is likely to be sustained as there is enough room for future development. Indeed, as of 2008 only 26 percent of the world population were online.[19] More importantly, in China, the emerging centre of economic power, roughly 22 percent of the population has access to the Internet.[20] In India, another rising power, only 5 percent of the population is on the net.[21] In the wireless sector, a proxy of the future traffic growth could be seen in sales of smartphones. In 2013, the total shipment of smartphones is projected to be 289 million, a substantial growth from 54 million in 2005.[22]

Actual traffic predictions from Cisco confirm this picture: wired traffic will soar fourfold between 2009 and 2014.[23] Video applications will contribute to this growth to a large extent because the share of Internet video alone will increase from about 30 percent of consumer Internet traffic to about 57 percent in 2014.[24]

The booming traffic determines future capacity requirements. For example, from 2009 to 2015 transatlantic capacity should increase from about 15 Terabyte/s to more than 60 Terabyte/s, an increase of more than four times.[25] Similar predictions hold for the wireless segment. According to Cisco VNI Forecast “mobile data traffic will double every year through 2014, increasing 39 times between 2009 and 2014;” 66 percent of the mobile traffic will be accounted for by mobile video in 2014.[26]

---

18 See AT Kearney (2010).

19 See AT Kearney (2010).

20 On the basis of information from Google public data.

21 On the basis of information from Google public data.

22 See AT Kearney (2010).

23 Deutsche Telekom also predicts the growth of the peak traffic on its own wired net from about 1 Terabyte/s in 2010 to about 2 Terabyte/s in 2012 and to 10 Terabyte/s in 2015.

24 See Cisco VNI forecast. Internet video is defined as online TV or VoD that can be downloaded or streamed for the PC or TV screens. Other forms of video include video communications (video calling, monitoring and webcam) and P2P.

25 See Global Internet Geography, by TeleGeography research © 2009 PriMetrica, Inc.

26 Deutsche Telekom expects that 27 petabyte will be downloaded via its wireless network in 2012, a significant jump from 16 petabyte in 2010, representing an annual growth rate of roughly 70%.

Fact 2: Over the course of the day, traffic volumes fluctuate greatly and high levels of congestion might be reached.

Traffic consumption is distributed unevenly over the course of any given day. In Europe, peaks are being observed between 4:30 p.m. and 9:00 p.m. in 2010: according to the Sandvine's report, bandwidth utilization reaches the highest levels during this time on an average day of 2010. In off-peak periods bandwidth utilization falls dramatically and lingers around one-fifth of peak capacity utilization between 1:00 a.m. and 5:00 a.m.[27]

The high volatility of traffic implies that either there is congestion during peak periods or alternatively a very large amount of capacity sits idle during off-peak periods – in Europe this would imply that significantly more than 30 percent (and up to significantly over 80%) of capacity is idle for most of the off-peak hours.

Fact 3: New applications require high-quality transmission standards.

A number of new applications (e.g., 3DHD video, cloud gaming, video conferencing) require a particularly high quality of transmission. The quality could be measured with a number of metrics, for example, delay, jitter, speed and packet loss, to list the most important ones.[28] However, all the measures are eventually determined by the total amount of bandwidth (i.e., the maximum number of bits of information that could simultaneously be sent along a wire), the quality of traffic management on the net and the priority given to a particular data package.

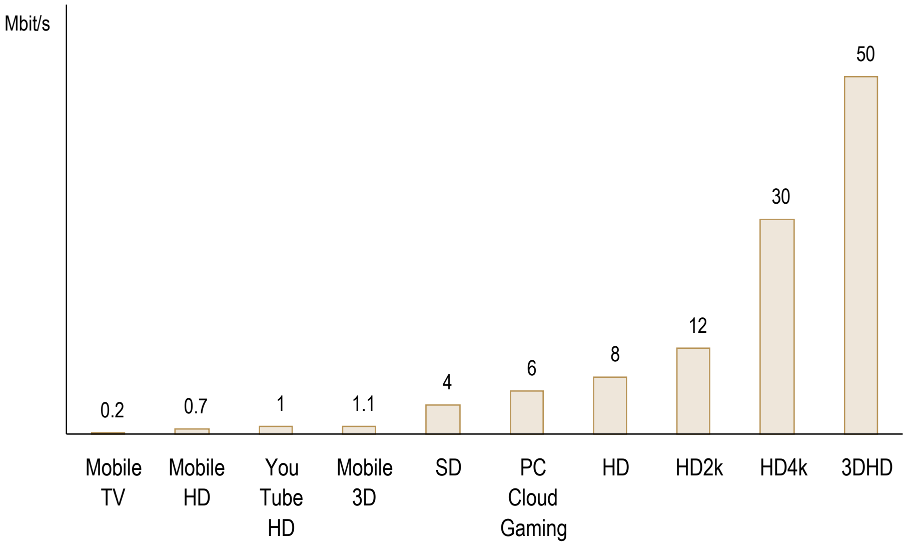
The graph below shows requirements of several new applications with respect to the speed of packet delivery. For example, streaming a YouTube video in HD quality requires 1.1 Mbit/s of transmission speed while streaming a 3D video in HD quality needs 50 Mbit/s.

---

27 See Sandvine (2010). The peak-period traffic surge is mostly accounted for by web browsing and real-time entertainment.

28 *Delay (latency),* which is essentially the time a bit of information, sent by an application, waits in a queue somewhere on the net, is obviously determined by the privileges given to the bit and by the broadness of the "pipe." *Jitter* is just volatility of delays between subsequently delivered bits, which together form a piece of content, such as streaming a clip on YouTube. *Speed* is an amount of bits, which arrive to a user's computer or are transmitted together along a wire each second; it is obviously capped by the amount of the bandwidth in the wire; it is also influenced by delays in the previous parts of the net. *Packets,* bunches of related bits, could be *lost* because of signal imperfections (mainly relates to the wireless transmission); also a packet, being considered time-elastic, could be dropped in times of severe congestions to be reloaded from the content provider again later; a packet could also be dropped if, because of delays it experienced, it becomes outdated for the requesting user. Thus, the packet loss measure is again determined by the amount of bandwidth and privileges the packets were granted.

Figure 3: Requirements of applications with respect to speed



Source: Deutsche Telekom.

"Content delivery networks" (CDNs), such as Akamai and Limelight, provide quality of service enhancements to certain content providers based on specialized technologies. As explained previously, these third parties store and serve content from a set of strategically placed local servers. This reduces the loads on the Internet backbone as well as the backhaul (aggregation networks), increases the reliability through content duplication and advanced load balancing algorithms and in general improves quality of data transmission for end users. Although, the data between a CDN and an end user is handled on a non-prioritized, "best effort" basis, the technology ultimately improves the end user's experience. Moreover, an increasing number of large content providers (such as Google) provide these services internally by creating a global network of local data centers ("server farms").

These facts indicate that some content providers are willing to pay premiums to CDNs for their services. This provides evidence of existing demand for the services providing quality exceeding that of a best effort network, even in the environment in which ISPs cannot differentiate. According to Buyya, Pathan and Vakali (2008), revenues of CDNs specializing in video content will increase from below US$300 million in 2007 to over US$1.4 billion in 2012 representing an annual growth rate of 36 percent.

Fact 4: End consumers are currently priced so that they experience little or no incentive to control the traffic they generate.

On the one hand, all end consumers typically pay flat rates, which might be differentiated with respect to the maximum available bandwidth. This is, for example, reflected by the OECD recent Global Communications Outlook which states that "broadband also remains largely a flat-rate subscription in most countries."[29] The flat rate implies that the end users' traffic consumption is unlimited.[30] On the other hand, end consumers significantly differ in their patterns of Internet usage. Time Warner Cable, an Internet service provider with 8.4 million broadband customers, reckons that the "top 25 percent of users consumed 100 times more data than the bottom 25 percent of users."[31] In a similar vein, Deutsche Telekom states in its response to the EU public consultation that 3 percent of their mobile customers generated 53 percent of the IP traffic in 2009.[32]

Essentially, this means that the current flat rate system of billing is subsidizing large consumers to the detriment of moderate users. It also implies that some very moderate customers, those who only need to check e-mails once per day, may be priced out of the market. These inefficiencies and re-distributional consequences cannot be circumvented as long as the market offers only fairly undifferentiated flat rates.

Fact 5: Peer-to-peer applications might jeopardize the payment balance under traditional transit agreements.

P2P technology partially circumvents transit via lower tier providers, thus reducing transit payments by content providers. At the same time, overall traffic is not reduced significantly. As P2P applications have gained importance in recent years, the amount content providers pay under transit agreements might no longer be a good approximation of the costs they produce on the entire network. To

---

29 See OECD (2009).

30 However, in some countries ISPs implement data caps which limit the amount of traffic that can be downloaded per month. For example, all surveyed Canadian and Australian ISPs, 82% of Belgium and 40% of British ISPs impose a data cap (OECD data available at http://www.oecd.org/dataoecd/22/46/39575020.xls). Nevertheless, the data caps are set at high levels, which hardly restrict the actual consumption (the Canadian cap is set at 76,727 Mb and the British one at 10,500 Mb per month).

31 See Lowry (2009).

32 See DT response, page 4 (http://ec.europa.eu/information_society/policy/ecomm/doc/library/public_consult/net_neutrality/comments/01operators_isps/dtag.pdf).

corroborate, in 2008 the peer-to-peer file sharing accounted for 32 percent of the total traffic on the Internet and for 22 percent of the global downstream traffic.[33]

P2P applications are predicted to lose relative importance in the overall Internet traffic. Indeed, in 2009 the share of peer-to-peer in the total global traffic dropped to 20 percent (from 32% in 2008) and the share in the downstream global traffic decreased to 16 percent (from 22% in 2008).[34] Moreover in 2010 the share of the peer-to-peer in the downstream traffic was only 8 percent in Europe and 13 percent North America.[35] However, some providers of traffic-intensive applications like video/radio broadcasting or videoconferencing have started to exploit the P2P approach for their services, most prominently Skype. The tendency could reverse the declining trend in the future.[36]

Fact 6: Network management practices allow a more cost-effective way to satisfy demand than over-provisioning.

Internet content and applications are becoming increasingly diverse in their network-performance requirements. New services like medical telemetry, network gaming, and video streaming need much higher quality of service than traditional content like e-mails and news. That necessitates additional investments from the side of ISP. However, under a net neutrality regime, additional capacity built to facilitate the new services must inevitably become available for the traditional traffic. In other words, if an ISP is set to enhance the quality of video streaming on its net, it needs not only to make enough investments to increase the quality of the video but also to bring quality of all other content to the same level.

Clearly, such undifferentiated treatment of content makes any quality-enhancing initiatives on the Internet much more expensive. Houle et al (2007) compare network investments required to facilitate the quality requirements of new services in two regimes: in one regime ISPs are allowed to differentiate the quality between two different types of traffic; in the other all traffic has to be treated equally (on a best effort level). The study finds that, to provide the same level of quality to new and traditional applications, ISPs would need to invest 60 percent more into infrastructure capacity under the net neutrality regime. To put it simply,

---

33 Sandvine (2009), Global broadband phenomena.

34 Sandvine (2009), Global broadband phenomena.

35 Sandvine (2009), Global broadband phenomena.

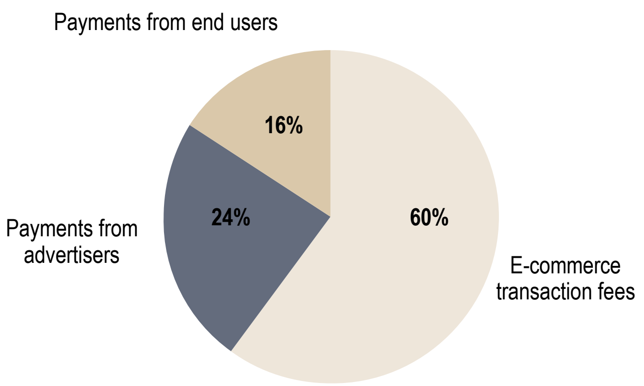
36 Skype is now the world's largest long-distance telephone provider with a market-share of more than 13%. Source: TeleGeography (http://www.telegeography.com/cu/article.php?article_id=31718).

the net neutrality regime makes high-quality experience of streaming video and online gaming significantly more expensive for consumers.

Fact 7: Content providers earn the largest share of the overall revenue in the Internet value chain.

Content providers earn money on the Internet by delivering advertising, by charging commissions on transactions taking place on e-commerce platforms and by selling content and services to end users. According to a study by AT Kearney, revenue earned by content providers in relation to non-business consumers amounted to €165 billion in 2008. Out of that amount, the largest part (€99 billion) came from merchants, paying for their placement on e-commerce platforms. Advertising revenues accounted for €40 billion. The largest part (72%) of the advertising revenues was earned by search engines and consumer publishing services. Viewers contributed only €26 billion through subscription or transaction-based prices.[37] Figure 4 below illustrates the distribution of the total revenue earned on the Internet.

Figure 4: Distribution of total revenue earned through non-business users by content providers, 2008



Source: AT Kearney (2010).

Furthermore, content providers grab the largest share of the revenue earned on the Internet. In 2008, 62 percent of the total revenue[38] was earned by content and

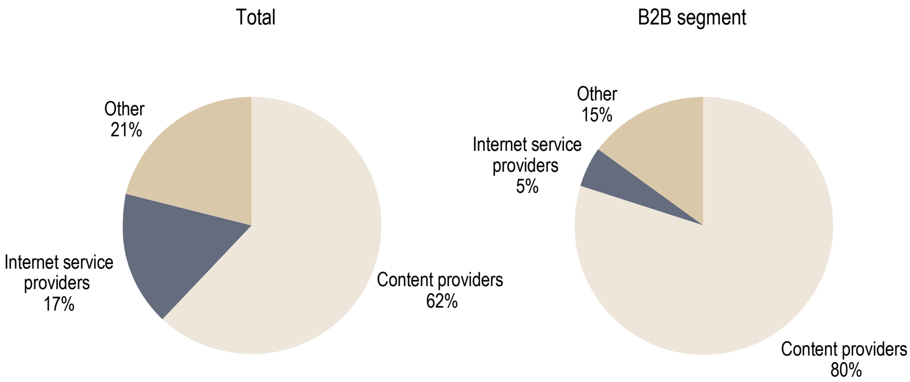
---

[37] See AT Kearney (2010). The revenue numbers were converted from US$ using the exchange rate US$1.47 per €1.

[38] The total revenue includes money earned by content providers and Internet service providers as well as content owners (TimeWarner, EMI, BBC), providers of enabling technology and services (Akamai, PayPal and DoubleClick) and user interface providers (Firefox, Symantec, and Apple).

service providers, while Internet service providers cashed only 17 percent. In the B2B segment of the industry[39] the revenue discrepancy was even larger. Content providers appropriated 80 percent of the total revenue, while Internet service providers only counted for a paltry five percent (Figure 5).[40]

Figure 5: Distribution of total Internet revenue, 2008



Source: AT Kearney (2010).

Looking at this revenue split between content providers and ISPs retrospectively, we observe that on the one hand revenues earned by content providers have gone up. For example, online advertising spending in Europe rose from €7.2 billion in 2006 to €14.7 billion in 2009.[41] Similar dynamics can be observed for online advertising in the US. On the other hand, prices for international and national transit as well as for Internet access for end users and for transit prices charged by the Internet service providers have declined. Indeed, in major cities like Hong Kong, London and New York, transit prices decreased continuously from 2005 to 2009.[42] This downward trend is also found in (national) transit prices in major cities around the world.[43] These plunging prices have to be seen in connection with

---

39 By the B2B segment of the industry we understand those Internet service providers and content service providers as well as content owners, providers of enabling technology and services and user interface providers, who serve business customers.

40 See AT Kearney (2010).

41 Information according to Interactive Advertising Bureau (IAB).

42 For example, the price in Hong Kong dropped from above 60 US$ Mbit/s in 2005 to 30 US$ Mbit/s in the first quarter of 2009 (Global Internet Geography, by TeleGeography research © 2009 PriMetrica, Inc.)

43 Prices in major cities in Europe and North America have declined with an average compound rate of around 20% from 2005 to 2009, cities in Asia by 10 to 35% and Latin American cities by 30 to 40% (Global Internet Geography, by TeleGeography research © 2009 PriMetrica, Inc).

several factors, such as increases in the number of autonomous systems, as well as the expansion of capacity and overall traffic and the accompanying expected decrease in unit costs. The TeleGeography annual survey finds that global capacity deployment has exceeded 60 percent since 2007. Peak as well as average traffic on international connections has increased on similarly high scales with some variety between regions. Furthermore, average prices declined in (nearly) all reported OECD countries from 2005 to 2008.[44]

A revenue forecast suggests that the disproportional revenue distribution will become even more pronounced in the future. While revenues of Internet service providers serving non-business customers are to increase by 6 percent, revenues of content providers serving the same customer group are expected to grow 16 percent between 2008 and 2013.[45] For example, revenues of IPTV operators, who will generate most of the traffic in the future together with other video service providers, will climb from US$17.5 billion in 2010 to US$46 billion in 2014 representing an annual growth rate of 27 percent.[46]

Moreover, investors consider the business of Internet service providers substantially less attractive than that of content providers. Indeed, share prices of prominent Internet service providers have hardly increased since 2004, while share prices of prominent content providers have soared by more than 200 percent since that time – despite the financial crisis in 2009.[47] This evaluation by the stock market makes it much more difficult for Internet service providers to raise capital for investments in net capacity.

Fact 8: The segment of content providers is becoming increasingly concentrated.

The Internet is becoming a more and more concentrated economic system with a relatively small number of participants (hosting, cloud and content providers) accounting for the increasing share of the total traffic. According to Internet Observatory 2009 Annual Report “out of the 40,000 routed end sites in the

---

[44] The average compound growth rate varies form +5% in Germany to -46% in the Czech Republic whereas speed increased significantly over the same time period in most countries. See OECD broadband portal (http://www.oecd.org/document/54/0,3343,en_2649_34225_38690102_1_1_1_1,00.html#prices). With regard to the mobile sector, Ofcom data shows that, while mobile traffic increased from the last quarter of 2007 to the last quarter of 2009 by 2300%, revenues of mobile operators barely doubled in the same time span.

[45] See AT Kearney (2010).

[46] See MRG, Inc. (2010).

[47] See AT Kearney (2010).

Internet, 30 large companies – 'hyper giants' like Limelight, Facebook, Google, Microsoft and YouTube – now generate and consume a disproportionate 30 percent of all Internet traffic." From the report it also follows that in 2007 thousands of Autonomous System Numbers (ASN) contributed 50 percent of the traffic, while in 2009 only 150 sites account for the same amount of the traffic.[48]

With an increasingly concentrated content provider side, it can be expected that the share of the jointly generated surplus that ISPs can appropriate is going to deteriorate. This would support the revenue forecasts whereby the revenue split between ISPs and content providers is going to shift even more toward content providers.

48 According to Compete, a web analytics company, the top 10 websites accounted for 31%of US page views in 2001, 40% in 2006 and about 75% in 2010. See http://www.compete.com/.

3. Typology of possible future business models

A business model covers the commercial relation between the ISP and all relevant players in the market. It also includes the level of vertical integration that the ISP is envisaging as this in turn determines the type of commercial relationship between different partners in terms of supplier or competitor. This report analyzes potential business models from the point of view of an eyeball ISP.[49] This section first determines the dimensions that business models can potentially cover. It then anchors the current business model within those identified dimensions. In a second part, it examines which characteristics of the identified dimensions constitute potentially important elements of future business models of eyeball ISPs. The last part offers a description of a number of different business models which could emerge or become more prominent in the future.

3.1. Dimensions of new business models

The dimensions of a new business model relate to the quality and pricing structure for each commercially relevant partner and the level of vertical integration of the eyeball ISP. The following table provides an overview of the general dimensions and characteristics of a new business model from the perspective of an eyeball ISP.

[49] When identifying and evaluating different business models it is of importance who actually implements the business model - an ISP those primary assets are the content it carries (content ISPs) or an ISP those primary assets are the users they are connecting (eyeball ISPs). This differentiation affects in particular the direction of traffic, the relative bargaining power in bilateral negotiations and the type of current interconnection agreements between the different players.

Table 2: Overview over dimensions of business models

Commercial partners	Quality scheme	Pricing scheme	Vertical integration
Content provider ▪ Directly vs. indirectly connected CP ▪ Small, midsized, large ▪ Including infrastructure based content aggregators	One uniform quality class ▪ Relative vs. absolute standard ▪ Static vs. dynamic representation	One uniform flat rate ▪ With or without volume cap	Traffic transmission and server integration
ISPs ▪ Eyeball vs. content ISP ▪ Access vs. transit ISP	Differentiated quality classes ▪ Relative vs. absolute standard ▪ Number of quality classes	Volume-based transaction prices ▪ Peak pricing ▪ Two-part tariffs	Infrastructure and content integration
End user ▪ Commercial vs. non-commercial ▪ Heavy usage vs. standard usage	Exclusive quality classes	Differentiation ▪ Quality classes ▪ Customers ▪ Services ▪ In-/outbound traffic	Additional third-party services

Source: ESMT CA.

The relevant **commercial partners** for eyeball ISPs can be classified in three broad groups: (1) content providers, (2) ISPs and (3) end customers. Within those three groups of players it might be necessary to distinguish further subgroups as their commercial relationship with eyeball ISPs are significantly different. For content providers it might be relevant to distinguish between those content providers which are directly connected to the eyeball ISP and those which are linked through other content ISPs or infrastructural content aggregators. Indirectly connected content providers do not have direct contractual relationships with the eyeball ISP relationship and, consequently, the business models which are feasible for those partners might be fairly restricted. Furthermore, business models can be distinguished according to small, mid-sized and large content providers including (infrastructure-based) content aggregator providing hosting as well as fast delivery services for several content providers as a subcategory of large content providers.[50] The size of a content provider might influence its bargaining power. Additionally, the administrative costs of implementing certain business models might depend significantly on whether it is dealing with a few large content providers or a myriad of smaller players.

In relation to ISPs, other eyeball and content ISPs can be distinguished.[51] This distinction is of importance – among other factors – as eyeball ISPs face higher costs than content ISPs due to the more expensive provision of last-mile access to end customers in comparison to business customers. Furthermore, traffic between an eyeball ISP and another eyeball ISP is normally more balanced than between an eyeball ISP and a content ISP. Furthermore, it might be useful to keep in mind that some ISPs might focus on providing long distance transmission capacities and might be less integrated into the last or first mile. Such transit ISPs naturally have less access to direct funds from users on either end of the Internet than access ISPs and essentially would rely on transit payments. However, at the same time it should be kept in mind that the major players – in particular tier 1 operators – are usually integrated to a significant degree into first- as well as last-mile provision.

End users might be businesses or private persons.[52] Furthermore, (private) end users can be distinguished with respect to their usage intensity.

---

50 Infrastructure-based content aggregators provide infrastructure-based services to content providers. Thus, in the classification of the commercially relevant partners they could be seen as either among content providers stressing the aggregation function or as a form of other ISPs stressing their involvement in infrastructure provision. Since content aggregators are, however, only to a limited degree integrated into the transmission infrastructure, at this point they are subsumed under the role of content providers.

51 Some economist might argue that the business relationship between ISPs and content providers does not require any direct relationship between the ISPs and content providers. Rather, ISPs may differentiate their business models with respect to content providers while these charges are being passed on through various ISPs in order to ultimately face content providers (cascading payments). They would go on to argue that within the current net neutrality debate, the role of agreements between ISPs is limited. In this report, however, we analyze the effects of different new business models. In this respect, it might be worthwhile to keep this group of players as a separate contracting segment as the filtering of content through other networks might restrict the feasibility of certain business models. Furthermore, depending on the market situation on various layers of the Internet some aspects like double mark-ups or cross-financing of content providers through end users might be of importance.

52 The political debate on net neutrality focuses on the potentially detrimental effects on private customers as their access to information might be hampered through traffic management techniques.

Turning to the **quality dimension**, one needs to carefully distinguish whether one is talking about quality in terms of the first or last mile of the connection or whether it is about the quality of transmission within the realm of the shared Internet resource.[53] The distinction is of importance as, for the first and last mile, content providers and users essentially do not share the resource.[54] In contrast, once the traffic reaches the network of networks, it has to share the available capacity with traffic which originates from other users. An analogy to the car traffic system might illustrate this point: whereas your private driveway gets you access to the road system and you are free to use it any time, you have to share the motorway with many other cars. The net neutrality debate focuses on the question of how to allocate the shared resource to different users. In contrast, there is no question of quality differentiation around the first or last mile bandwidth which is current practice today. The following thus focuses on the quality of transmission within the shared resource.

The quality dimension can have several characteristics, such as uniform quality classes, differentiated quality classes as well as exclusive quality classes for some players. Quality classes could either be defined in absolute or relative terms. An absolute standard would guarantee a minimum quality which implies that a given network structure can only guarantee access to a maximum number of customers at any given time.[55] In contrast, relative standards define quality with respect to lower quality classes. So instead of guaranteeing, for example, a certain absolute speed, a higher quality class would offer relatively faster transmission.[56] An obvious difference between absolute and relative quality classes lies in the fact that the latter does not necessarily need adjustment over the course of time as the Internet infrastructure evolves. The specific way quality classes are defined can have

---

[53] Quality can technically be defined in many different dimensions (e.g., bandwidth, delay, jitter, packet loss). We abstract from the technical dimension and treat quality as a black box.

[54] Of course, on the first and last mile there might also be some degree of sharing of the capacity. However, the closer traffic gets into the core of the network, the more it has to share the available resources with other traffic.

[55] Greatly simplifying the structure of the Internet, absolute quality classes could be illustrated with differently sized (end-to-end) tubes.

[56] If, for example, the high quality class offers twice as fast a transmission compared to the lower quality class, this would imply that the ratio of transmission speeds between the two classes in time of no congestion would be (roughly) identical to the ratio in times of congestion. Another way of thinking of different relative quality classes is rooted in the queuing theory: instead of keeping the transmission speed roughly proportionate over time between different quality classes, one could think of different relative qualities as forming two different queues at congested nodes where items in the high-quality queue would have to be delivered ahead of items in the low-quality queue.

implications in terms of the incentives to degrade quality or to invest in infrastructure capacity. Furthermore, it can be of importance in the regulatory assessment of the different business models. Generally, the various effects of quality differentiation will be discussed with a general notion of quality differentiation in mind; however, where necessary, the specific ways of implementing quality differentiation will be considered.

A complicating factor concerning the implementation of traffic with different qualities is that both sides of the traffic, that is. the sending as well as the receiving end, could have different quality agreements with the ISP. In this case, the business model needs to define a clear matching rule between the sender's and the receiver's quality classes.

Another interesting aspect within this framework lies in the number of different quality classes. If we abstract from the presence of multiple networks, it appears that the focus on two quality classes captures the essence of the effects which can be introduced within the (non-exclusive) quality differentiation approach. However, as soon as this level of abstraction is lifted, new interesting aspects in relation to the standardization process between multiple networks might arise.

A business model incorporates different **pricing structures** for different commercial partners. Generally, in terms of pricing, an ISP has a whole range of possibilities from flat rates to transaction-based prices and from uniform prices for all traffic/services to differentiated prices in many dimensions. Differentiation can occur with respect to the commercial partners or quality classes. As discussed above, it is obvious that due to – among other things – the different needs for Internet usage, offers to end users will have a different structure than offers to content providers (as is currently the case). Additionally, differentiation with respect to different content providers is conceivable. Such differentiation could relate to the individual content provider's (or end user's) identity as well as to groups of content providers and end users.

As a last dimension to Internet business models stands the potential to **vertically integrate** into various aspects of content provision. This opens a virtually unlimited space of possibilities for ISPs and different ISPs have already moved forward on different elements. Some prominent examples of vertical integration include: (1) server landscapes, (2) online content platforms, (3) third-party services such as billing and (4) targeted advertising.

Before analyzing the various elements of business models, the following tries to anchor the current status quo within the discussed dimensions. Table 4 gives an overview of the concepts with respect to content providers, interconnecting ISPs and end users. The table can be read from row to row.

Table 3: Current business model (from an eyeball ISP's perspective)

Partners	Quality scheme	Pricing scheme	Vertical integration
Content providers	Uniform quality class (differentiation only in access quality based on first-mile bandwidth)	Small CPs: typically flat rates (differentiated for first-mile bandwidth) Midsized CPs: typically volume-related pricing Large CPs: either transit or peering agreements (see below)	Various levels of vertical integration (depending on eyeball ISP), in particular in online platforms and caching activities
ISPs	Uniform quality class	Typically peering agreements with transit ISPs such as AT&T, large content providers e.g., Google and content aggregators such as Akamai Transit agreements for smaller ISPs, smaller content providers and smaller content aggregators: traffic-related pricing (in terms of measured peak bandwidth utilization)	N/A
End user	Uniform quality class (differentiation only in access quality based on last-mile bandwidth)	Flat rates (differentiated for last-mile bandwidth)	N/A

Source: ESMT CA.

Abstracting from the modest levels of traffic prioritization going on today, the current system is characterized by a uniform transmission quality for content providers, other ISPs and end users. In terms of pricing schemes we have to distinguish between small, mid-sized and large content providers. Small content providers who are less sophisticated in predicting the levels of traffic generation tend to connect with flat rates. In contrast, mid-sized and large content providers face an array of different contracting options which typically all involve some level of volume-related charges. Interconnecting ISPs can be distinguished by the type of price contract they have with the eyeball ISP: either peering or transit. End users, on the other hand, in general hold flat rate contracts. Furthermore, currently ISPs are vertically integrated into content provision to a varying degree. For example, in comparison to some US operators like Comcast, the level of content provision of the German operator Deutsche Telekom is fairly limited. Vertical integration into the activity of content aggregators also differs depending on the ISP: most ISPs are not operating within this area to a visible extent; however, there are some exceptions to this rule, such as AT&T.

3.2. Essential elements of future business models

The previous section illustrates that the space for alternative business models is large. Essentially, each dimension can be crossed with all other dimensions in order to generate potential business models. This section identifies elements of viable business models by analyzing them from the point of an ISP's profit maximization. Thus, this section tries to answer the question of whether a certain pricing or quality strategy is likely to increase profits or revenues of the ISP. Often the answer to the question depends on certain characteristics of the market situation, in particular the existence of market power, transaction costs and the behavioral responses of consumers.

When discussing future business models, one should keep in mind the question why ISPs have not established the proposed new business models already. For this, it is important to realize the fundamental changes which have occurred in the past. Those changes are evident in the facts displayed in section 2.3. At this point, we would like to stress again the following: increase in traffic volumes (see fact 1), currently high congestion during peak times (see fact 2), increased demand for higher quality levels (see fact 3) and a shift in relative cross-group externalities with more revenue being made on the Internet (see fact 7).

However, even in a changing environment it might be difficult to implement one-dimensional changes. For example, it might be difficult to change the pricing scheme without offering a value proposition in return, that is, higher quality classes. To the extent that such feedback effects are relevant, it is essential to consider not one-dimensional, but multi-dimensional deviations from the status quo.

3.2.1. Viable quality schemes

Quality differentiation tries to achieve a more efficient allocation of traffic during peak times by giving priorities to certain types of traffic. Economic theory predicts that product or quality differentiation increases total welfare. In many circumstances, this implies that ISPs' profits also increase (see Economic first principles in section **4.1** for more detail). Thus, in general, quality differentiation has the potential to increase ISPs' profits and should thus be analyzed as an element of a viable business model.

Of course, as previously explained, quality differentiation can come in different shapes. Among the many different possibilities, the two stylized strategies described in the following can be considered as focal points. The first strategy would implement a dynamic best effort model for traditional services and a new flexible model for innovative services. The second model would implement a number of quality classes that content providers can choose from.

The first strategy, henceforth called "*Best Effort Plus,*" is inspired by the US debate on net neutrality regulation. There, the discussion currently evolves around a model where the best effort network as we know it would be preserved for traditional services and truly innovative services would enjoy some regulatory holidays.[57] Thus, one more or less realistic future scenario for quality differentiation could be that there exists a uniform quality class of best effort for all "traditional" applications on fixed infrastructure. However, ISPs would have the ability to develop and market innovative new online services (such as e-health, or new entertainment and gaming options) outside the best effort class as long as those services do not interfere with the "continued development of Internet access services" (best effort).[58] Consequently, those novel services would be marketed to end users independently of the general best effort Internet.

Of course the definition of what constitutes a novel service would be crucial in order to ensure that innovative new online services are not used to circumvent the non-discrimination rule for traditional services. When discussing the effects of such quality differentiation, it is assumed that a clear definition of what constitutes a traditional service and what constitutes an innovative new online service can be achieved.

The definition of quality outside the best effort network would be bilaterally negotiated between the ISPs and the content providers of novel services. In this setup, eyeball ISPs could offer individual and exclusive quality levels to developers of new services. However, it should be kept in mind that only a limited number of

---

57 The FCC has summarized the issues on which there appears to be consensus among stakeholders: First, ISPs should not be able to block lawful content. Second, there should be some form of protection against discrimination of content. Third, ISPs should be allowed to engage in reasonable network management as long as they are transparent about their network management practices. Whether or not specialized or managed services should be excluded from the regulation on broadband Internet access services and how to deal with mobile broadband access still appear to be open issues and are broadly discussed. See http://www.fcc.gov/Daily_Releases/Daily_Business/2010/db0901/DA-10-1667A1.pdf.

58 This approach is for example also described in the joint proposal of Google and Verizon, see http://googlepublicpolicy.blogspot.com/search/label/Net%20Neutrality.

individual and exclusive quality levels can technically be provided for. Thus, in this future scenario, either a number of content providers of novel services use a limited number of different quality levels non-exclusively or a few providers have exclusive agreements.

In the second strategy, a business model might emerge where content providers or end users can choose among a number of pre-defined quality classes. We refer to this model as "*Quality Classes*." Different quality classes would be defined on top of the current best effort level and every content provider or end user would have the ability to choose among different quality classes. Since all current and future content providers including those which are not directly connected to the eyeball ISP would be able to choose from different quality levels, it is likely that a form of standardization or interoperability process would emerge.

This strategy is inspired by the current ongoing industry activities – in particular, also within the mobile infrastructure industry – to define standards for different quality classes. According to industry sources, a number of ISP associations, such as, for example, the GSM Association, are already in the process of trying to define a number of different quality classes for the mobile segment.[59] Due to the fragmented nature of the Internet, it appears that there are currently a number of independent standardization efforts. One major challenge of such a business model would be getting a sufficient number of relevant players to agree on certain standards. It appears that to date no single solution as to the number and definition of quality classes has emerged from the standardization efforts.

Whereas the content provider is likely to invest in higher quality classes for individual applications, the end user typically consumes a variety of different contents. Each of those contents might need a different quality of transmission in order to optimize the quality of experience for the end user. Thus, ideally, the consumer would purchase different quality levels for different types of contents and services that she consumes. However, if ISPs do offer different quality levels for different services to end users, this could potentially be a very confusing and complex marketing exercise. In contrast, the decision problem for the content provider is significantly simpler as he would choose a certain quality level for each of its services. Thus, in a scenario where end users pay for higher quality classes, they are going to be confronted with flat rates for quality classes that are optimized for a certain content consumption pattern.

---

[59] See for example http://gsmworld.com/our-work/programmes-and-initiatives/ip-networking/ipx_pci_trials.htm. According this information, the negotiations on the mobile industry (GSM Association) resulted in a system called IP Exchange which is currently in "Pre-Commercial Implementation Trials."

The two described future quality scenarios share some similarities but differ also in a number of aspects. Table 5 summarizes the differences. Those two scenarios together with the uniform quality best effort standard of the status quo form the basis of the analyzed business models.

Table 4: Differences between quality scenarios

	“Best Effort Plus”	**“Quality Classes”**
Access to different qualities	Unlimited access to best effort Limited access to higher qualities (only novel services)	Unlimited access to best effort Unlimited access to higher qualities
Exclusivity	Non-exclusive best effort Potentially exclusive higher quality classes	Non-exclusive best effort and higher quality classes
Standardization/interoperability	Not at the core of the business model	Integral part of the business model

Source: ESMT CA.

3.2.2. Viable pricing schemes for content providers and end users

The following section discusses each pricing option from zero pricing to two-part tariffs and from no differentiation at all to differentiation on the basis of the commercial partner’s identity.[60] Table 6 gives an overview of the pricing options for different end users which are considered within this section.

60 The following discussion focuses on (medium-sized) content providers which are directly connected to ISPs and currently pay according to traffic generation. Smaller content providers who are currently on flat rates constitute a marginal group in terms of revenue and profit generation. They are thus not explicitly treated here. If transaction costs are sufficiently small, smaller content providers can be offered the same pricing schemes as larger content providers. Very large content providers who are also vertically integrated into parts of the infrastructure provision tend to have peering agreements. Those content providers are going to be subsumed under the discussion on interconnecting ISPs.

Table 5: Pricing schemes for different commercial partners of eyeball ISPs

Pricing scheme	Directly connected content providers	End user
Zero pricing	Dismissed	Dismissed
Flat rate versus volume-based transaction prices (Two-part tariffs)	Currently: 95th percentile volume in Mbps * price (EUR/Mbps) Introduction of fixed-fee element: dismissed	Currently: flat rates for different services Introduction of data caps: viable Introduction of volume-based pricing: dismissed
Peak pricing	Viable	Dismissed
Differentiation between customer groups	Financing mode: viable Type of service: connected to qualities Size of CP: connected to volume discounts	Heavy and light users: viable Business and private: connected to qualities
Differentiation between in/outbound traffic	Dismissed	Dismissed
Differentiation on the basis of identity	Partly viable	Dismissed

Source: ESMT CA.

These pricing options are analyzed from the perspective of the functions they fulfill. Thus, the section is structured into the following subsections:

- **Price levels between different market sides**: This part deals with the question of the optimal sharing of the financing burden of the infrastructure among the different sides of the market. It discusses, in particular, zero-pricing schemes.
- **Structure of prices within one market side**: The structure of prices within one market side determines on the one hand participation and on the other hand the actual usage of the platform. This subsection thus primarily discusses two-part tariffs.
- **Congestion aspects**: Volume-based pricing in contrast to flat rates introduces traffic-steering elements. Furthermore, stricter congestion-based pricing might ameliorate congestion problems on the existing infrastructure.
- **Revenue sharing aspects:** This part discusses termination charges and related revenue sharing agreements.

- **Differentiation aspects**: This part describes potential methods to differentiate prices between different players in the market.

The discussion of those pricing options partly depends on the market structure which is assumed: typically in the economics literature, the ISP industry is viewed as a two-sided market where the ISP balances the demand from the two sides taking the respective externalities that one group exerts on the other group into account. However, one could also view the market not as a two-sided market, but as a vertical supply chain where the transmission of data is simply an input to content providers when selling their products to end users (one-sided market perspective).

Taking a vertical perspective on the industry, the ISP's service loses part of its platform character: one can think of the content provider having a direct commercial relationship with the end user and thereby being able to (partially) levy any additional fees by the ISP onto the end user. This would imply that the structure of the ISP's pricing (or the level of the prices) of the two market sides becomes of lesser importance as content providers can simply pass additional fees through to end users. For most content providers which currently generate sizable revenues on the Internet this vertical view of the industry appears accurate to at least some extent.[61]

Consequently, we are going to look at both perspectives when analyzing pricing strategies.

3.2.2.1. Price levels between different market sides

The level of prices of the different players essentially determines which side of the market (end users or content providers) shoulders the larger share of the costs of infrastructure investment. In this context, it is illustrative to discuss zero-pricing regimes as those clearly allocate the burden of infrastructure investment costs on one side of the market. In a one-sided market structure, charging one customer group zero prices would certainly not reflect cost structures. It would thus be

---

[61] Take, for example, online selling platforms such as eBay or Amazon. They charge the seller of products a commission once a successful business has occurred. If the seller is able to differentiate price according to an online and offline search procedure, it can simply increase the online price by the amount of the commission. Thereby, the structure of the pricing between the two market sides becomes less important. Another example would be online search platforms like Google which finance their efforts through targeted advertisement. Here the transmission mechanism is less straightforward than in the eBay example as the online search platform would only be able to increase the number of advertisements, increasing the burden for consumers, in response to an increase in the fee by the ISP.

difficult to rationalize within a well-functioning market. However, within the perspective of a two-sided market, zero pricing of one market side could occur independently of the competitive situation in the market. In fact, zero pricing on one side in order to increase participation of that side is a feature often observed in two-sided markets. As examples, one can think of shopping malls or credit cards.[62]

In the context of the ISP industry, zero pricing for **all content providers** does not seem to be a feasible option. It would imply that all the financing of the infrastructure is burdened on the shoulders of end users. As end users become more and more important to content providers in terms of revenue generation – for example, advertising-driven content provision relies on the end users as eyeballs for the advertisement, e-commerce platforms rely on their ability to match sellers and buyers – it seems implausible that an efficient pricing structure involves end users entirely financing the infrastructure.[63] However, zero pricing for **all end users** also does not seem to be a feasible option for the near future, although one could imagine that end users become so important to content providers that content providers would be willing to subsidize the access of end users. In the current political debate over net neutrality, proponents of the regulatory intervention have argued that increasing prices for the content provider will lead to inefficiently low incentives to innovate for them. Within this debate it appears to be unrealistic that a business model involving zero prices for end users financed by higher prices to content providers would be feasible.

Thus, none of the stylized business models considers zero pricing for any of the two market sides.[64]

---

[62] There are essentially two views on the market structure which have different implications for price levels: (1) essentially one-sided markets with negligible cross group externalities or (2) essentially two-sided markets with significant cross group externalities. Typically, pricing in one-sided markets is impacted in particular by marginal cost functions as well as the elasticity of firm-specific demand. In the special case of perfectly competitive markets, pricing reflects marginal costs. In contrast, in two-sided markets prices additionally reflect the relative level of externalities that one side of the market exerts on the other side. One observes a so-called waterbed effect whereby the optimal pricing implies that the participation of one market side is being “subsidized” by higher prices on the other market side. See principle 4 in section 4.1.

[63] In terms of the economic literature on two-sided markets, the observed trend would imply that the cross group externalities from consumers to content providers are becoming more important.

[64] However, a new business model could also set the monetary price for end users to zero and finance the infrastructure provision by targeted advertising. This cannot be considered a zero-pricing model as end users pay in terms of nuisance costs from

3.2.2.2. Structure of prices within one market side: Two-part tariffs

This section addresses the question of the optimal structure of prices on one side of the market from the point of view of the ISP. It discusses the question to what extent the pricing should rely on fixed and to what extent on variable elements (two-part tariffs). On an intuitive level the fixed component of a two-part tariff could be seen as regulating the participation of content providers whereas the volume-based component could be attributed a traffic steering function (see also details on common-pool characteristics in section **4.1** on economic first principles). However, general economic theory on two-sided markets has not been able to make clear predictions on the optimal pricing structure when two-part tariffs are possible and the platform is not a monopolist. An array of different two-part tariffs can emerge in oligopoly.[65] This multiplicity of equilibria makes predictions about the tariff structure difficult.

On the most general level, Armstrong (2006) puts forward that if platforms can coordinate on the pricing structure with the highest profit for the platform, then a pricing structure stressing the transaction-based element is likely to emerge. The intuition is that with fixed fees only the competition between platforms for different members of one side of the market is fierce as an additional member of one side increases the valuation of the platform to all members of the other side. This cross-group externality can be partly internalized by using transaction-related pricing. This reduces the competitive pressure as a platform does not need to worry as much about how well it is performing on the other side of the market when payment is proportional to successful interaction. This argument provides some guidance as to the desirability of flat rates: it would imply that it is likely that flat rates are less profitable than two-part tariffs including a transaction-based element.

Thus, for **end users** who are currently priced with flat rates it appears that two-part tariffs might be a superior option. The introduction of transaction-based prices is likely to decrease the fixed component in comparison to a flat rate, potentially increasing the number of end users connecting to the Internet.[66] In the

---

advertising, although (targeted) advertisement might also fulfill some information and signaling functions, see Bagwell (2007) or Cabral (2008).

65 See Rochet and Tirole (2006), Armstrong (2006), or Weyl (2010).

66 A recent paper by Samanta and Pan (2009) models the ISP industry as two-sided markets with the purpose of analyzing the optimal pricing structure. They take into account that end users are the ones initiating traffic. When calibrating their model with demand elasticities for end users, they find that a negative fixed fee might be optimal for a monopoly ISP. Another paper by Bedre-Defolie and Calvano (2010) provides another

telecommunications business, even negative fixed components are not a novel feature as, for example, in mobile telephony handhelds are often subsidized by the mobile operator. However, in the past, consumers have rejected the volume-based component in connection to the Internet. One potential reason might be that in comparison to mobile telephony it is much more difficult for consumers to exactly track the volumes they generate by using different online services. Returning to a volume-based model might therefore be difficult and not a feasible business option for ISPs. Instead, an ISP might introduce a simplified version of volume pricing which imposes strict data caps in flat rates. This is discussed in more detail in section 3.2.2.5.[67]

With respect to **content providers** it should be noted that pricing is currently strictly volume-related and does not include a fixed element. Thus, the introduction of two-part tariffs would likely imply that the fixed fee would be negative and the volume-based prices would increase. This might be a feasible option. However, it appears at odds with the current discussions on net neutrality where some ISPs have announced that they intend to increase prices to content providers generally in order to receive more of the increasing revenues that content providers make online.[68] Generally, we consider that by introducing two-part tariffs to content providers, there is only limited room for ISPs to increase profits within the framework of two-sided markets.

One could also assume the perspective of a one-sided market. Economists typically discuss the benefits of two-part tariffs in vertical industries where at a low level of the supply chain there exists market power: an input supplier with some level of market power which charges simple volume prices will set those prices above marginal costs, which in turn implies that the final good is sold at a price above marginal cost, introducing inefficiency. In comparison, if the input supplier charges in terms of a two-part tariff, it will optimally set its volume-related price to

modeling approach (applied to credit cards) which introduces exogenous shocks at the usage stage: membership and usage decisions are distinct because they are made at different information sets. The model shows that the platform sets inefficiently low transaction prices to the side which determines the extent of trade at a given number of participants while over-taxing the other side. Both papers provide some insight that the current pricing structure is inefficient from a social welfare perspective.

67 Volume caps for end users are a form of volume discounts. Volume discounts for content providers have not been discussed explicitly as they are only partially interesting in the given setting where content providers do not fine-tune the amount of traffic they are generating. More importantly, the general debate around pricing issues hinges rather on the fact that existing capacity is overused, thus a stimulation of volume generation through volume discounts does not appear to be the goal.

68 See http://www.netzwelt.de/news/81792-telefonica-google-netznutzung-zahlen.html.

marginal costs and extract the monopoly rent through the fixed fee. This way, the inefficiency within the supply chain is reduced and the profits of the input supplier increased. In the present context, it appears that the level of competition of ISPs with respect to content providers is high.[69] Thus, there seems to be only a limited concern about the inefficiency within the system due to the exertion of the market power of content ISPs. Due to this, introducing two-part tariffs for content providers – although a feasible option for content ISPs – does not appear to be at the heart of the question on future business models.

Thus, we expect only limited profit potential from two-part tariffs or the introduction of a fixed-fee element for content providers. That is not to say that we are not going to see two-part tariff structures in the future, but merely that two-part tariffs for content providers do not seem to lie at the heart of future new business models.

Thus overall, none of the stylized business models focuses on two-part tariffs.

3.2.2.3. Congestion aspects

The recent past is characterized by congestion during peak hours; traffic forecasts predict that this problem (on the given infrastructure) might become more prominent. Thus, a more pronounced peak-load pricing element might adapt traffic pricing to these new circumstances. Since the infrastructure would be used more efficiently, it is likely that the ISP has an incentive to introduce such pricing.

Once the participation decisions of content providers and end users are taken, the bulk of the traffic – at least in the classical server-client architecture of the Internet – will be initiated through the end user. Due to flat rates, end users currently have no incentives to limit their traffic in congested times. As long as volume-based pricing for end users is not feasible or can only provide for insufficient incentives to use the existent level of infrastructure efficiently, it appears to be particularly important to keep a volume-based element for content providers. Even though content providers do not initiate the traffic generation, they influence traffic generation through the design of their content as they can create volume intensive and less intensive applications. This is, in particular, of importance as the end users are often unaware of the amount of traffic they are generating with different applications.

---

69 For example Ofcom states that: "*However, as discussed above, concerns about consumer detriment would arise only if the ISP engaging in such practice had market power but we do not currently consider that this is the case in the UK.*" (see Ofcom (2010), para 4.19).

The current volume pricing is aligned with the 95th percentile of generated traffic during a month. This type of volume-based pricing reflects "congestion" on an overall level. It does not take into account at what time the traffic of a particular content provider is generated as it charges the same price for traffic generation during peak hours in the early evening as during the remaining off-peak hours.[70] Congestion depends on the timing of the traffic volume in combination with the consumption pattern of all other traffic generators. Thus, merely counting volume is a poor measure for congestion.

Instead of mere volume-based pricing, an ISP could charge different prices depending on the time of the day as another proxy for congestion-based pricing.[71] A peak-load pricing scheme[72] would implement this idea by using the described 95th percentile method for 5-minute samples in different periods over the day (potentially varying by geographic area). The respective price per Mbps would adapt with the expected (average) level of congestion during the different periods. However, peak-load pricing is an imperfect implementation of congestion pricing. As such it imposes inefficiencies with respect to a perfect implementation. For example, with a limited number of defined pricing periods, usage patterns might circumvent the high-price classes by generating traffic close to those high-price periods (instead of spreading traffic more evenly in off-peak periods). This in turn would imply that congestion problems during peak periods might be ameliorated at the cost of intensified congestion around the peak periods. Furthermore, introducing complex pricing systems like peak-load pricing can also induce additional transaction and administration costs; in particular content providers

---

70 The current pricing scheme according to the 95th percentile of generated volume is also sometimes referred to as peak-traffic pricing. However, it should not be confounded with congestion-based or peak-load pricing which takes into account the level of congestion imposed by traffic on other users and thus usually includes a time of the day component.

71 Although theoretically it is possible to implement pure congestion pricing, this is practically impossible: local prices would be changing every second depending on the local level of congestion taking into account alternative routing possibilities and the level of congestion on those routes. Even if content providers would accept this type of highly flexible pricing, it is not clear whether the decentralized nature of the Internet would permit the collection of the relevant information in order to be able to implement the system.

72 Typically, peak-load pricing is applied for non-storable commodities whose demand fluctuates periodically. In such circumstances, uniform prices lead to the provisioning of capacity able to meet demand at the peak. Consequently, in off-peak periods the capacity is significantly under-utilized. However, since capacity comes at a cost, peak-load pricing addresses the pricing inefficiencies which result from uniform prices. The general idea of peak-load pricing is that demand in peak periods should pay for the extra costs of capacity provision due to the demand peaks. For a survey on peak-load pricing literature, see Crew, Fernando and Kleindorfer (1995).

could incur costs for keeping track of price changes in order to be able to adjust their behavior accordingly. When judging peak-load systems, it is thus necessary to compare the benefits from decreased congestion with the additional transaction costs.[73]

Although theoretically some form of peak-load pricing can be implemented for the end user as much as for the content provider, the implementation of efficient peak-load pricing for consumers appears to be more difficult. Abstracting from transaction costs, the idea of peak-load pricing would be to introduce as many geographic/time period segments as possible in order to increase the efficiency of the system. However, private consumers are likely less willing to accept an increasing number of pricing classes than professional users such as content providers. Many different classes would constitute a very complex optimization exercise for the end users. In fact, experimental results in local telephony services have indicated that consumers are unlikely to accept any pricing scheme involving more than three periods over the course of the day.[74] Thus, it is likely that an implementation of peak-load pricing for content providers could be a more effective tool for tackling congestion. Still, it is conceivable that some form of peak-load pricing (in terms of different flat rate options differentiating bandwidth with respect to peak and off-peak hours) could also be implemented with end users.[75]

Summarizing, we expect that stronger elements of congestion-based pricing could ameliorate the current congestion issues. We thus present a business model which focuses on this issue. This business model combines a uniform quality level with a peak-load pricing scheme for content providers. For end users, the model

---

73 In the past, there have been some experiments around the implementation of peak-load pricing in local telephone services. The empirical evidence from those experiments suggests that the benefits in terms of decreases in congestion resulting from the implementation of those peak-load pricing systems did not exceed (or only to a very limited extent) the associated implementation costs. Furthermore, the results suggest that peak-load pricing can at most support three pricing classes with respect to end consumers. This clearly limits the benefits of peak-load pricing as the number of pricing periods is confined to be small. However, this result refers to end consumer pricing and is thus only a poor indicator for pricing professional users such as content providers. For a short overview on this empirical literature, see Yoo (2008).

74 See Yoo (2008), 37.

75 Westnet and some other carriers have put offers online that could allow for bandwidth differentiation during peak and off-peak periods, however, so far no differentiation has been established, see http://www.westnet.com.au/broadband/plans.html. In particular, peak load pricing is currently discussed for the mobile Internet access as an alternative to data caps. See for example http://www.wired.com/epicenter/2010/07/peak-data-hours/.

introduces some volume-based element through differentiated volume caps in flat rates, but does not go as far as introducing peak-load pricing systems. The business model does not appear at the core of the future business choices – given the political discussion on net neutrality – it is thus considered more as a benchmark or a slightly amended status quo.[76]

3.2.2.4. Revenue sharing aspects

Revenue sharing aspects of pricing between ISPs and content providers relate to what is known as termination charges in the economic literature on two-sided markets. Termination charges apply to a setting where end users are single homers, that is, they are only connected to a single platform, and content providers who want to reach all available consumers are present on all the platforms through multi-homing. Currently, the interoperability and the fact that no content provider is blocked from eyeball networks implies that content providers pay their access fee to a single access ISP and are thereby connected to all end users. If, however, in the future eyeball ISPs employ a business model whereby they determine the number of content providers to which their customers have access on the web, they might decide to employ revenue sharing mechanisms or termination fees.[77]

A non-differentiated revenue sharing rule could specify that eyeball ISPs get a percentage of the revenues that accrue to the content provider due to the access to the eyeballs of this specific ISP. Every time a user accesses a content provider's services and the content provider receives some revenue from this, this revenue could be split between the content provider and the ISP as complementary input provider. This proposal is not new in the online industry as currently in particular e-commerce facilitators employ a similar pricing model whereby for every successful e-commerce transaction the seller pays a commission to the facilitator.

Often revenue sharing might however also incorporate a differentiation aspect. For example, one could imagine a system of transaction charges which is differentiated according to the financing mode of the content provider. Currently, we are aware of three different financing models:

---

[76] Of course, this type of pricing might also be part of any other of the business models. However, the stylized models try to focus on different elements of new models and separate out the effect.

[77] Please note that the general discussion of two-sided market pricing in the section on two-part tariffs does not distinguish between single and multi-homing strategies. In fact, the choice of single or multi-homing is to some extent also endogenous to the pricing system as platforms can change their pricing schedule so that they encourage or discourage single homing.

- **Advertising**: In the advertising model, revenues accrue per click (e.g., online search engines). Thus, a revenue sharing agreement between the ISP and the content provider would specify a price per click that originated from the eyeball ISP's network.
- **Subscription**: In the subscription model, revenue is generated by end users who subscribe for a certain period of time for a specific (unlimited) service (e.g., online gaming or dating platforms). Here a revenue sharing model would imply that the content provider pays a (monthly) fee for every subscriber who gets access over the eyeball ISP.
- **Transaction**: Finally, a content provider business model which relies on revenues from individual transactions (music or video downloads, e-books) could share this revenue on a transaction basis.

Furthermore, ISPs could try to implement individualized prices or revenue sharing agreements for -at least a subset of - content providers. In this scenario, any type of revenue sharing could emerge as the result of individual bargaining. Furthermore, such an approach might lead to exclusive content agreements on the one hand and to the exclusion of certain content on the other hand. In other words, individualized price negotiations could result in a closed platform approach as partially witnessed in the mobile market.

Those different revenue sharing options might be an interesting way to go for ISPs. Thus, it will be considered an integral part within one of the business models. In order to be able to implement a revenue sharing agreement, it is likely that the ISP will have to offer value added in return in many bargaining situations. For this reason, we only consider the revenue sharing option in business models which offer higher qualities. Revenue sharing is thus a pricing alternative for the "Best Effort Plus" model. As a counter-scenario, the "Quality Classes" model focuses on offering differentiated qualities on a non-discriminatory basis and abstracts from the possibility of revenue sharing agreements.[78]

3.2.2.5. Differentiation aspects

Generally, economic theory predicts that in a variety of market situations price differentiation increases total welfare (for more detail, see section **4.1** on economic first principles). It therefore appears reasonable to look at future business models which differentiate prices according to different customer groups.

---

78 Business models incorporating different quality classes might also offer revenue sharing agreements based on different types of content providers. The above distinction is chosen for expositional purposes.

Before turning to a discussion of which types of price differentiation are feasible, the conditions for being able to price discriminate are discussed. Price differentiation among directly connected content providers or end users is only feasible if the ISP is able to refuse to contract with the commercial partner. This seems to be a fairly innocuous condition for end users. However, if an ISP refuses to contract with a particular content provider, this content provider can simply access the Internet via a different access provider. Thus, in order to be able to price discriminate among directly connected content providers, the ISP also needs to be able/allowed to block content on their networks coming from other interconnecting networks.

The relevant question is then, along which lines (apart from quality) does differentiation appear sensible? For **content providers**, the following differentiation dimensions have been considered:

- **Delay sensitivity of application**: This type of third-degree price differentiation is closely related to the question of quality differentiation and is not discussed separately.
- **Size of content provider**: To the extent that volume discounts are present to date, second-degree price differentiation according to size already exists. However, third-degree price differentiation would price users of different sizes with different unit prices independent of generated traffic. This could be an interesting pricing option if the content provider's valuation for units of volume changes with size. This would be the case if, for example, online search engines had a systematically higher valuation of every bit transported than video download services. This, of course, is an empirical issue. However, there do not appear to be any industry insights hinting at such a correlation between size and valuation. Furthermore, it is difficult to disentangle third- and second-degree price discrimination with respect to size/volume. Thus, we will not consider this type of third-degree price differentiation.
- **Up- to downloading activity**: This type of differentiation would differentiate prices between content providers whose ratios of up- to downloading differ. To the extent that a different ratio of up- to downloading activity among content providers gives an indication as to the valuation of content providers for transmission services, it could be worthwhile implementing such a simple pricing scheme. One could argue, for example, that content providers who rely on sharing technologies tend to have a high ratio of up- to downloads. Simultaneously, those content providers tend to be less often for profit organizations. Thus,

their willingness to pay might be lower than others' whose ratio of up- to downloads is lower. However, it appears that there is no established empirical relationship along those lines and that industry insights do not argue in this direction. Thus, similarly to differentiation according to size, we do not consider this type of differentiation a viable business model.

Similarly, for **end users**, the following differentiation dimensions have been identified:

- **Heavy and light users**: Most of the traffic is generated by a limited number of end users. But heavy users currently pay the same amount as light users. Light users thus cross-subsidize heavy users.[79] Imposing volume data caps would restrict the level of cross-subsidy by light users. Furthermore, flat rates might be differentiated with respect to the data cap. This would increase the ability of ISPs to differentiate pricing between heavy and light users and introduce a volume-based element into the pricing scheme vis-à-vis end users. This differentiation would help increase participation as low-volume users enjoy lower tariffs and would thus be more willing to join the platform service. In the following, introducing data caps is considered a viable business option.[80]
- **Different types of services**: Another pricing scheme for private end users would involve a decomposition of the uniform flat rate into several flat rates for different types of services exhibiting specific characteristics with respect to sensitivity to congestion (such as VoIP, IPTV and Internet

---

79 This abstracts from the fact that heavy users might be more likely to choose more expensive flat rates with higher bandwidth allowances than light users. However, even if this is accounted for – as long as there is not a perfect match between heavy users and high bandwidth and light users and low bandwidth – there remains an element of cross-subsidy between heavy and light users.

80 In a similar fashion, ISPs can differentiate between heavy and light users by introducing different sets of two-part tariffs: heavy users tend to be attracted by high fixed components and a small volume based rate whereas light users tend to prefer low fixed fees (potentially negative) with relatively high volume-based rates. However, end users seem reluctant to the idea of volume pricing and this option is thus not considered. One way of implementing data caps would be to introduce soft caps according to some form of fair usage policy where consumers who consistently overuse their allocation can be forced to switch to a higher price class. This process would involve that the ISP is able to credibly threaten that it will refuse to deal with the end users if the end user does not switch to a higher price class with a higher volume cap. This process would enable users to learn more about their usage pattern in order for them to form a judgment on the appropriate flat rate for them.

access services). The implementation of this requires that the eyeball ISP is able to block certain types of specialized services for the end user or that certain services are not available on the (otherwise) open Internet. This type of decomposition of the flat rate is already present in the market in the sense that currently eyeball ISPs offer the following three flat rate options: (1) for Internet access service, (2) for Internet access services and fixed telephony and (3) for Internet access services, fixed telephony and TV packages (triple play). Although those services are currently not uniformly applying the Internet protocol, they are all running over the same (wire) infrastructure. Furthermore, it can be expected that in the future all those services will switch to routing data packets over IP due to the associated higher efficiency.[81] If the technical difference between voice or TV signals vanishes, voice and TV services are not easily distinguishable by the transmission technology from any other special service. Thus, it is conceivable that flat rates for other special services such as video conferences might arise as well. This pricing scheme is going to be considered under the "Best Effort Plus" model.

- **Business and private users**: A differentiation with respect to business and private users would be based on the fact that business users are expected to have a higher willingness to pay for higher quality classes than private users. For business clients Internet traffic can be regarded as a general input into the production process. Thus, faster transmission might have positive spillover effects within their businesses. This implies that business customers might display a higher willingness to pay for the prioritization of traffic. Thus, this type of differentiation is again closely linked to the question of quality differentiation and is thus only considered in conjunction with this, but not as an independent scenario.

In summary, only differentiation on the end user side between light and heavy users seems a viable business model element. This differentiation can be implemented via the introduction of various volume caps. This type of end user volume pricing is being considered in the modified status quo which does not offer

---

81 For voice telephony, there are currently two competing transmission mechanisms. Voice signals can be transmitted in the old-fashioned way as voice signals or can be transformed into data packets which are routed via the Internet protocol. In the future, it is likely that due to higher efficiency of routing data packets this mechanism is going to prevail. With respect to TV signals, the trend appears to also shift from traditional signal transmission onto transmission over IP.

differentiated quality classes but instead focuses on a better congestion management.[82]

3.2.3. Pricing with interconnecting ISPs

Generally, all proposed pricing schemes for content providers could also be applied to the interconnection agreements with other ISPs. However, it appears likely that a number of factors constrain the pricing behavior of an eyeball ISP with respect to other interconnecting ISPs. This section thus explores which pricing elements might be difficult to implement from the perspective of an eyeball ISP.

Generally, the theoretical literature analyzing the pricing incentives of ISPs in two-sided markets abstracts from the fact that the platform is split into many interoperable networks. However, as long as one assumes that the contracting options for the interconnecting ISPs are not limited, it is likely that the general qualitative results of two-sided market literature still persist. It can, however, be questioned whether the contracting options are not limited. Thus, this section tries to understand the fundamental issues related to the bilateral negotiations between the interconnecting ISPs in a multi-layered representation of the Internet.[83] It is illustrated that the inability to distinguish traffic of different content providers coming through the same content ISP might effectively limit the contractual options.

In the bargaining situation between interconnecting ISPs, the assets that content ISPs bring to the bargaining table are precisely the types of content that they have managed to contract whereas the assets of eyeball ISPs lie in the number and nature of their eyeballs. Under the assumption that content ISPs fiercely compete in order to be able to provide access to content providers, it could be expected that they offer their service at marginal costs. However, once a content ISP has managed to contract an attractive content provider this is going to give it bargaining power vis-à-vis the (monopolistic) eyeball ISP. Thus, first of all it should be noted that, despite the fact that content ISPs deliver access in a competitive situation, they cannot be considered as price takers, but may still bargain with eyeball ISPs: Efficient bargaining would then result in a situation where content ISPs manage to appropriate the cross-group externalities exerted by their content providers so as to pass the financial benefits arising thereof (at least partly) on to

---

[82] Of course, this type of pricing might also be part of any other of the business models. However, the stylized models try to focus on different elements of new models and separate out the effect.

[83] It therewith also tries to shed more light on whether the current situation of content and eyeball ISPs conserves the features of a two-sided market.

them. Thus, essentially we would expect to see that for attractive content the access provision is cross-financed by end users (and vice versa). This would in turn result in an access price for those contents which are below marginal costs.

If however one takes into account that content ISPs are bundling the content of many different content providers with different characteristics, the situation might change. As previously mentioned, different types of content might exhibit different cross-group externalities. For example, end users might be particularly sensitive to new gaming options whereas another new search engine would generate a lower utility for end users. In a one-layered Internet service provision, the ISP might find it profitable – and in fact this might also be socially optimal – to price discriminate content providers according to the relative strengths of their cross-group externality.[84] In order to be able to perfectly mimic a market situation with a single platform incorporating content and eyeballs, the eyeball ISP would need to be able to perfectly distinguish traffic stemming from different (types of) content providers at the point of interconnection. This would be essential in order to determine fixed and volume-related fees for the content providers. However, the content ISP has a function of bundling traffic. Thus, it is not obvious that traffic of different content providers could be or can be (technically) distinguished.[85] Furthermore such an implementation of differentiated prices might come along with significant transaction costs such as installing additional counters for traffic stemming from different types of content providers. Also, there might be historical reasons why traffic is not distinguished between different content providers despite the fact that in the current situation it might be efficient to do so.[86]

The inability to distinguish between different types of content might offer one reason for why content ISPs, despite being fairly competitive, could have hindered efficient pricing at the interconnection point. This might be explained with the following example: Suppose historically a content and an eyeball ISP negotiated a peering agreement. Now, the content ISP managed to contract with a few very attractive content providers. This might imply that their bargaining power is so

---

[84] For a detailed depiction of price discrimination in two-sided markets, see Weyl (2010).

[85] For example, Schuett (2010) sets the discussion of net neutrality in context to the fact that ISPs were not able to distinguish the origin of data packets in the past, but would be able to do so by now.

[86] For example it could be that historically content providers did not differ extensively with respect to their cross-group externalities. Furthermore, it could be that historically a clear cut distinction between content and eyeball ISPs did not exist. Instead, every ISP would have had more or less the same type of traffic generating members. In this case, efficient contracting between interconnecting ISPs could well have implied very simple pricing schedules.

strong that they manage to keep the peering agreement despite the fact that for the rest of the content the direction of cross-group externalities would imply positive transaction prices.

In consequence the inability to distinguish traffic from different content providers at the interconnection point – be it for technical, administrative or historical reasons – might restrict the contracting space and therewith result in different outcomes in the two- (or multi-) layered set-up in comparison to the one-layer set up.[87]

The reasons why it is difficult to change peering agreements are discussed above: if the peering partner has very valuable content providers, the eyeball ISP risks losing this content. This is in particular an issue with tier 1 peering partners. A tier 1 provider can be defined as an operator which can provide access to the entire Internet by only peering with other operators. In consequence, content which is only directly connected to tier 1 operators is not accessible unless one has a peering agreement with the tier 1 provider. In contrast, content which is directly connected to an ISP which maintains transit as well as peering agreements with other interconnecting ISPs can be assessed through various different channels. Therefore, de-peering with a tier 1 peering partner would essentially mean that the eyeball ISP loses access to the directly connected content.[88] This is not the case for higher tier peering partners.

In summary, the above discussion affects the business models in the following sense (assuming that eyeball ISPs cannot block traffic at the interconnection point from specific content providers in the future):

- Cascading payments: The inability to block certain content also implies that eyeball ISPs will find it difficult to convince content providers to change their access provider to the eyeball ISP. The eyeball ISP simply lacks a value proposition. Consequently, we do not consider the direct contracting of eyeball ISPs with all content providers in any of the business models and remain within the framework of cascading payments.
- Peering: Current tier 1 peering partners will stay as peering partners for any type of best effort delivery. De-peering with tier 1 operators is unlikely to be an option given that this would mean that certain content cannot be accessed by the end users of the eyeball ISP. However, for higher quality delivery some transit-type agreement could be expected.

87 This argument does not take any issues related to double marginalization into account.

88 This applies only to the content providers which single home with tier 1 operators.

Thus, for all business models, peering is going to persist for the best effort level. For the "Best Effort Plus" and the "Quality classes" model, peering agreements can be transformed into transit agreements for the higher quality levels.

- Differentiation: Transit prices might be differentiated with respect to peak times or quality classes. Furthermore, to the extent that revenue sharing agreements in the provision of innovative services in the "Best Effort Plus" model involve the cooperation between interconnecting ISPs, some differentiation on the basis of content providers might take place.

3.2.4. Vertical integration

As previously pointed out, there are a vast variety of vertical integration options. In the introductory chapter, integration into the server landscape, online content platforms, third-party services and targeted advertising were mentioned. While all of those options might be valuable for ISPs, only those options which have a direct impact on the interaction with relevant commercial partners are being considered here. This excludes in particular aspects of vertical integration which have mere cost implications for the ISP.

For example, ISPs could move further into the provision of server landscapes and enter further into the field of infrastructure-based content aggregation. Basically, investment into content delivery networks has two effects: firstly, content delivery networks allow delivering traffic with higher qualities. Secondly, ISPs can reduce the amount of traffic routed within their system holding the volume of delivered traffic to the end users constant. With sufficiently low investment costs into content delivery networks, an ISP might thus prefer to invest in this type of infrastructure over an overall expansion of the transmission capacity. In other words, faced with a need to double the capacity of its network, a system of "smart server caches" might achieve a similar effect at a lower cost. In consequence, ISPs might decide to invest in content delivery networks independent of the question of whether they offer differentiated qualities of services. However, the incentive to invest in content delivery networks is likely to be stronger if ISPs decide to offer different qualities. This vertical strategy is thus not further examined. According to a study by Ovum (2010), some operators such as AT&T have recently pushed into this market already.[89] The study acknowledges in particular that ISPs might be very well-positioned to enter this market as they might benefit from deep financial pockets, strong brands and extensive local networks. This strategy focuses on

---

89 Ovum (2010).

traffic steering and delivery and does not aim primarily to affect the bargaining position vis-à-vis relevant players, therefore it is not considered here.[90]

Another element of a vertical integration strategy could hinge on building a strong **online content platform** integrating own content as well as third party content. Many of the major ISPs have already got a platform presence online. The question would be to what extent this platform is pushed in the future, whether an open or closed approach is adopted and which other services are made available within the platform. Within the context of an online content platform, but also independently of it, several other **third-party services** such as billing or standardized technology interfaces could be offered. As an example in this area the recent acquisition of "ClickandBuy," an online payment provider, by Deutsche Telekom can be mentioned.[91] Here, in particular the question of which player holds the access to the end user in terms of billing might be crucial for the split of the surplus between content providers and ISPs. To a large extent this depends on whether consumers' trust is more with the eyeball ISP or with the content provider. Furthermore, ISPs might also have some cost advantages due to economies of scale. Those factors directly affect the bargaining outcome between content providers and the ISP and are thus considered as a relevant vertical strategy within this context.

In many classical examples of two-sided markets, end users are not priced involving financial payment streams (or not only), but instead they pay in terms of the nuisance costs of **advertisement**. This is, for example, the case in newspapers (online and offline), TV programming, online platforms such as search engines and social networks or offline search platforms such as Yellow Pages. Similarly, ISPs could propose a new business model with respect to end users that offers subsidized Internet access on the basis of accepting targeted advertisement in return. This would imply that ISPs offer new services to advertisers - a new customer group. This would also mean that the two intertwined two-sided markets, online search engines which coordinate advertisers and eyeballs on the one hand and Internet access platforms coordinating search engines and eyeballs on the other hand, converge further. However, this type of targeting of advertisement might breach consumer privacy directives and is not considered any further within this report.[92]

---

90 Notwithstanding the fact that entering the market of content delivery networks will of course affect the relationship of an ISP with current content delivery networks.

91 See press release at http://www.telekom.com/dtag/cms/content/dt/de/838220.

92 See for example the UK case on Phorm http://ec.europa.eu/unitedkingdom/press/press_releases/2009/pr0928_en.htm.

Billing technologies and other third-party services might help to gain bargaining power and increase the value proposition of the ISP which in turn allows for transferring content providers into more efficient and profitable pricing schemes. This is of particular importance in the "Best Effort Plus" model where innovative services are marketed to the end customer, however, the provision of third-party services might also be of relevance – in particular for the smaller content providers – in the "Quality classes" model.

3.3. Conclusion on future business models

This section summarizes the arguments in relation to elements of viable business models by proposing four different complete business models. In each of these business models different elements described above receive different emphasis. Also the mutual reinforcement between these elements differs. It is thus worthwhile to reshuffle the elements and discuss the resulting business models in more depth.

Each one of the business models focuses on a different aspect: The first business model stresses the possibility to tackle congestion problems through congestion-based pricing. The second model preserves the traditional best effort network, but gives ISPs more leeway with innovative services. The third model stresses the perceived need of different applications for various degrees of quality of service. The last model, however, puts the focus on consumer choice for higher quality levels.

3.3.1. BM 0: Congestion-Based Model

The focus of this business model lies in tackling congestion issues. Consequently, in this business model no quality differentiation of any sort is introduced. Instead, end users are priced according to a stronger volume-based component by introducing flat rates with differentiated data caps. Furthermore, for directly connected content providers as well as transit ISP partners, it foresees that the current maximum volume pricing is applied for different pre-defined peak and off-peak periods. As mentioned previously, it is unlikely that peering partners will accept any type of positive prices as long as no new product is introduced.

Table 7 gives an overview of the model anchoring it within the three dimensions discussed earlier.

Table 6: Anchoring of "Congestion-Based Model" (BM 0)

Partners	Quality scheme	Pricing scheme	Vertical integration
Content provider	Uniform best effort	Peak-load pricing: 95th percentile method for different times of the day/days of the week	N/A
ISPs	Uniform best effort	Peering as in current status quo Transit agreements with peak-load pricing: 95th percentile method for different times of the day/days of the week	N/A
End user	Uniform best effort	Flat rates with differentiated volume data caps	N/A

Source: ESMT CA.

3.3.2. BM 1: Best Effort Plus

The business model "Best Effort Plus" preserves the best effort Internet, but additionally introduces priority lanes for clearly specified, innovative services. For the remainder of this study, innovative services are assumed to not function on a best effort level as they are particularly sensitive to different quality of service dimensions. It is further assumed that innovative services can be clearly distinguished from traditional services. In order to keep the analysis simple, it is assumed that at the intersection with other ISPs only best effort traffic is handled, in other words innovative services are connected directly to the eyeball ISP. This assumption abstracts from issues of multi-lateral bargaining which might be relevant when several ISPs cooperate in order to deliver innovative services.

Content providers are priced as in the status quo if they operate on the best effort level. In contrast, content providers with innovative services which are transmitted outside the best effort with an individually specified level of quality are priced according to the individual negotiations between the eyeball ISP and the content provider. For interconnecting ISPs nothing changes in comparison to the current status quo: only best effort traffic is being handed over at interconnection points. Of course some innovative services might be delivered over a number of interconnecting ISP (e.g., e-health application which connected US-based patients with European-based specialist doctors). Those interconnecting ISPs are likely to bargain together with the content provider. Efficient bargaining implies that we do not lose much generality if we assume that a single ISP is transmitting the innovative services. End users can purchase Internet access on a flat rate basis.

Additionally, they can get access to the innovative services on a subscription basis (either flat rate or transaction-based fees).

Innovative services require the guarantee of a specified level of transmission quality by the ISP. This eventually implies that there is a greater level of vertical cooperation between the ISP and the content provider, for example, in terms of research and development. Furthermore, the ISP might expand its services to certain innovative content providers by offering them billing functions. Given the direct access to end users via the best effort network, this might be attractive – in particular for smaller content providers – in order to build up a customer base.

Again, table 8 gives an overview of the model anchoring it within the three dimensions discussed earlier.

Table 7: Anchoring of "Best Effort Plus" (BM 1)

Partners	Quality scheme	Pricing scheme	Vertical integration
Content provider	Dynamic best effort model for traditional services Higher qualities for innovative services	Pricing as in status quo for best effort Individual contracting over access to higher qualities (differentiation between customer groups possible, differentiation on identity only for non-vertically integrated ISPs)	High level of vertical "cooperation" between innovative services and ISPs ISP might offer third-party services, e.g., billing services
ISPs	Dynamic best effort model for traditional services	Peering and transit agreements as in current status quo	N/A
End user	Dynamic best effort model for traditional services Higher qualities for innovative services	Flat rates as in current status quo Flat rates or transaction based rates for innovative services	N/A

Source: ESMT CA.

3.3.3. BM 2: Quality Classes: Content Pays

Instead of partitioning the Internet into an open best effort and a (closed) platform for innovative services, the "Quality Classes – Content Pays" business model proposes to offer different standardized or interoperable quality classes which are accessible to all content providers on a non-discriminatory basis. Quality classes would be tailored to the specific needs for quality of different types of applications. In the example of the business model below, four quality classes are distinguished: (1) interactive, (2) multimedia, (3) critical, and (4) best effort. It is assumed that content providers pay volume-related prices according to the quality class which they are choosing.

Content providers can choose the quality class from the offer of the ISP which is connecting it. The access ISP would then have to ensure that the specified quality transmission can be achieved on an end-to-end connection potentially including other interconnecting ISPs. This in turn implies that ISPs coordinate their quality classes or make an offer to transport traffic according to the quality classes of other operators. Former peering partners will still transport best effort traffic on a settlement-free basis. However, for higher quality traffic transmission volume-related prices just like in transit agreements are going to be implemented. Equally, transit agreements are going to differentiate transit prices according to the chosen quality class.

For the end user, nothing changes in terms of the pricing. However, its quality of experience will change depending on whether the applications run on a high- or a low-quality transmission class.

Lastly, in this business model, ISPs are going to pursue quality enhancements via two routes. On the one hand they are going to use stricter traffic management tools; on the other hand, they are going to decrease the overall traffic by engaging in caching activities.

The different elements of the “Quality Classes – Content Pays” model are summarized in table 9.

Table 8: Anchoring of "Quality Classes – Content Pays" (BM 2)

Partners	Quality scheme	Pricing scheme	Vertical integration
Content provider	Four differentiated quality classes: High qualities: Three higher quality classes (Interactive, multimedia and critical) Best effort	95th percentile method for different quality levels	N/A
ISPs	Quality classes determined by connected content providers Some sort of interoperability agreements between ISPs on quality classes	Transit partners: 95th percentile method for different quality levels Peering partners: 95th percentile method for three high-quality levels and peering for best effort	N/A
End user	No choice of quality classes	Flat rates as in current status quo	N/A

Source: ESMT CA.

3.3.4. BM 3: Quality Classes: User Pays

This business model is similar to the previous as ISPs offer different quality classes. However, in this scenario, ISPs offer the different quality classes to end users and not to content providers.[93] In this case quality classes would be devised which would match the different usage patterns of end users: end users who frequently use interactive applications might choose the quality class which is more apt at dealing with such applications, that is, that offer a particularly low level of delay and jitter in comparison to an offer that focuses on multimedia applications offering particularly low packet loss and high bandwidth.

[93] The distinction of who pays for the higher quality classes is made in order to separate the effects of the payment structure. It is, of course, conceivable that a mixture of the two business models BM 2 and BM 3 could arise in practice.

Table 9: Anchoring of "Quality Classes – User Pays" (BM 3)

Partners	Quality scheme	Pricing scheme	Vertical integration
Content provider	No choice of quality classes	Pricing as in status quo	N/A
ISPs	No choice of quality classes	Pricing as in status quo	N/A
End user	Four differentiated quality classes: High qualities: Three absolute quality classes (Interactive, multimedia and critical) Best effort	Flat rates for different quality classes	N/A

Source: ESMT CA.

4. Theoretical analysis of effects

This section compares the welfare effects of different business models to the status quo. In contrast to the previous section which focused on the profit prospects, it assumes the perspective of a social planner who evaluates different future scenarios.

Given the complexity of the Internet and the broadness and variations in business models – which may partially coexist with each other – a rigorous assessment within a unique theoretical framework is not feasible. Instead, we define a number of fairly general and robust results from the economic literature ("economic first principles") that are relevant for the assessment of the expected effects of new business models from the social welfare point of view. These economic first principles are put forward in the first part of this section.

In the second part of this section we carry out the analysis, identifying the major social benefits and costs linked to each business model.

Finally in part three of this section, we discuss the implications of various forms of net neutrality regulation – which is currently debated on the EU level – on those business models.[94] Net neutrality regulation, if and when formally implemented in some shape or form, has the potential to reallocate resources among industry participants, affect optimal pricing strategies and affect their investment and innovation incentives. Through these effects, the shape of net neutrality regulation is going to affect which business models are going to be at all feasible, which are going to thrive, and which are going to become obsolete. Therefore, when assessing the potential impact of net neutrality regulation, one needs to consider how the regulation may affect future business models and the costs and benefits associated with them. The last section thus discusses regulatory options in relation to the different business models.

Before proceeding with the analysis, we would like to point out that because theoretical results have often been achieved under a set of quite restrictive assumptions and because a net neutrality regulation would have impacted a number of variables in many different ways, the theory alone is of limited usefulness to predict the overall effect of a net neutrality regulation and an assessment including empirical components would likely be required. As a recent survey on the theoretical literature concluded:

---

[94] See DG Media, http://ec.europa.eu/information_society/policy/ecomm/library/public_consult/net_neutrality/index_en.htm.

> A general theme that emerges from the literature is that the welfare effects of net-neutrality regulation tend to be ambiguous. A zero-price rule may increase or decrease welfare, depending on the relative magnitudes of the network externalities between consumers and content providers, among other things. A non-discrimination rule may increase or decrease short-run welfare, depending on assumptions about the nature of competition between content providers, the organization of the sale of priority service, and various other parameters. Furthermore, a non-discrimination rule may increase or decrease investment in network capacity and thus long-run welfare.[95]

4.1. Economic first principles

In this section we summarize and characterize a number of fairly general and robust results from the economic literature that are relevant for the assessment of the expected effects of new business models from the social welfare point of view.

Principle 1: Common-pool resources are characterized by congestion and suboptimal level of investment

Common-pool resources are similar to pure public goods in the sense that there exists difficulty of developing physical or institutional means of excluding beneficiaries (so-called **non-excludability**). Because of that, common-pool resources are characterized by a strong temptation to free ride on the efforts of others, which leads to suboptimal investment in improving the resource. On the other hand, common-pool resources share with private goods the attribute that one person's consumption subtracts from the quantity available to others (so-called **rivalry**). Because of that, common-pool resources are subject to the problems of congestion or overuse unless use limits are devised and enforced.

Common pools can be described as a system comprising a (stock) resource system (e.g., the Internet infrastructure) and (flow) resource units (e.g., data packets or content files).[96] A certain level of the resource system can only support a certain amount of the resource units to be consumed by the users of the common pool. It is frequently the case that the resource system is jointly owned and operated, while the resource units are withdrawn from the system by private individuals (appropriators).[97] Devising property regimes that effectively allow sustainable use

95 See Schuett (2010), 11.

96 See for example Hess and Ostrom (2003).

97 Common-pool resources may be owned by national, regional or local governments, by communal groups, by private individuals or corporations or (as is the case for the Internet) used as open-access resources by whomever can gain access. The exact property

of a common-pool resource requires rules that limit access to the resource system, but also another set of rules that limits the amount, timing and technology used to withdraw the resource units from the resource system.

The term Internet is often used to describe a **physical network** (i.e., infrastructure consisting of optical fiber and copper wires, routers, switches, servers and end user workstations/computers interconnected with one another) which may be analyzed within a common-pool resource framework. However, the term Internet can also be used to describe **information resources** (i.e., content), that includes webpages, documents, images, databases, audio and video files, indexes, catalogues, and other resources accessible using this physical network infrastructure.[98]

The physical network and information resources are interdependent, but separate and materially different. Internet Service Providers make decisions about the provisioning of physical network capacity. Content providers affect the utilization of physical network capacity by making decisions about the application design and network protocols used which affect the volumes and rates of data transmissions over the Internet. Finally, the behavior of end users determines how the Internet is utilized. They decide which websites to visit, what software to run and download, how many messages and of what type and size to send and at what time of the day to perform these actions. Even though the impact of an individual user on the overall Internet infrastructure seems negligible, on aggregated coordinated (or simply correlated) the behavior of millions of users is significant in terms of consuming network capacity and creating congestion.

A number of fundamental design features and properties allow treating the Internet as a common-pool resource. First, the Internet is a distributed, **non-centralized system** with no single central management or owner. Second, the Internet can be thought of as a "**network of networks**," comprised of many autonomous but interconnected networks, which in the aggregate constitute the Internet.[99] Third, the Internet has no explicit hierarchy and communication between constituting networks is often based on **peering relationships** which assume symmetry between participating networks at the point of

---

regime governing the common-pool resource is to a large extent irrelevant, although each has its own sets of advantages and disadvantages. Examples exist of both successful and unsuccessful efforts by governments, communal groups, cooperatives, voluntary associations and private individuals or firms to govern and manage common-pool resources.

[98] See Bernbom (2000).

[99] An implication is that no individual member network can be easily excluded, since it is connected to the Internet at many access points simultaneously.

interconnection.[100] Fourth, the Internet is based on **open standards** which facilitate communication and interconnection but at the same time make exclusion more difficult. Finally, the **fundamental design principle** of the Internet, which posits that any data to be transmitted over the Internet is divided into small packets that are then sent toward its destination independently and possibly through multiple different routes, also makes exclusion difficult. In summary, all these features contribute to non-excludability as a basic feature of the Internet and a common-pool resource.[101]

The rivalry of the Internet manifests itself as congestion in the case of overuse of the physical network infrastructure. The Internet becomes congested when too much data is sent at the same time over the same part of the physical network.[102]

The problems of overuse and congestion translate into a problem of underinvestment in a dynamic setting. ISPs to an extent have incentives to expand their facilities to keep pace with increasing demand. Access providers will lose customers if they cannot offer reliable and high-quality access. Content providers will not attract users if their site is unreachable or slow to respond. All these incentives are sufficient to provide some level of investment, but are likely not sufficient to reach its socially optimal level given non-excludability and the resulting incentives to free ride on others' investment efforts.

A potential solution to the congestion problem of common-pool resources are **restrictions of access** to the resource through introduction of usage policies, including reasonable network management policies. Such policies are often informally implemented at a local level, for example, government-funded high-performance networks often implement acceptable use policies which restrict access to primary research and education uses.

---

[100] In contrast, exclusion with transit agreements would be much easier to achieve, e.g., by setting a sufficiently high price.

[101] See Bernbom (2000).

[102] Internet congestion may differ in terms of its intensity and duration. It may be transient when it is a result of a brief and unanticipated burst of data traffic on the network. Internet congestion may also be sustained, for example, if too many users simultaneously try to use the same resource, e.g., a website of a popular event. Finally, Internet congestion may become chronic if the aggregate demand for network capacity consistently exceeds the capacity available. Transient and unanticipated sustained congestion can be viewed primarily as resulting from the appropriation problem, while chronic and anticipated sustained congestion (if it occurs) can be seen as a provisioning problem. See Bernbom (2000).

Another solution and probably the most relevant to the net neutrality discussion requires additional incentives to invest in capacity, so that congestion can be kept low at the socially optimal level. Common-pool resource literature suggests that congestion problems can be resolved by **changing the appropriation rules** that govern the network resources. This includes development of new protocols, practices or methods to improve network performance or diminish the impact of certain high-demand uses of the network. One example would be the development of different data quality classes for different network applications with different characteristics in terms of their tolerance to latency or jitter, volume and rate of data they send or receive, or other important characteristics. Differential service would allow traffic to be separated depending on its characteristics and then transmitted at appropriate level of service necessary, instead of the current practice to send all traffic at the same "best effort" priority and competing for the same scarce network capacity.[103]

Principle 2: Product differentiation increases total welfare

Economic literature makes an important distinction between concepts of price discrimination and price differentiation. In essence, if different consumers are charged different prices for the same product, then price discrimination takes place. In contrast, differential pricing denotes a situation in which a firm charges a menu of prices for a number of products or services with different characteristics. We understand that some versions of net neutrality regulation would prohibit both price discrimination (ISPs charging different prices for the same service) and differential pricing (ISPs charging different prices to content providers for the provision of different levels of service). Because the two concepts are different we analyze them separately, focusing on product and price differentiation in this principle and analyzing the principle of price discrimination separately in the next section.

The introduction of product differentiation quite generally generates positive welfare effects. Broadly speaking, product differentiation increases welfare because it increases the number of available choices and allows heterogeneous consumers to choose consumption bundles more closely suited for their individual preferences. End users benefit from increased diversity, variety of content and delivery mechanisms and are able to better tailor the consumed bundles to their heterogeneous preferences. For end users, product differentiation (both in terms

---

[103] See Bernbom (2000). Moreover, congestion externalities can be also ameliorated through consolidation among interconnecting ISPs. Once the network of networks is more consolidated, it becomes easier to steer traffic through pricing mechanisms without discriminating access between formerly "separated" networks.

of content availability and quality classes for delivery) is beneficial because it increases their chance of finding a good that more accurately meets their needs.[104]

More specifically, abstaining from strict net neutrality restrictions in the future might allow ISPs to offer differentiated quality of service enhancements. These optional products, unavailable under strict net neutrality regulation, would be better tailored to some content provider needs, for example, offering reduced latency for latency-sensitive applications. Product differentiation with respect to end users is already part of the industry landscape, with most ISPs offering multiple subscriptions differing with respect to maximum download/upload speeds or additional services such as telephony or TV, but in the absence of net neutrality regulation the number of available choices could increase even more in the future.

For content providers, the multiple optional quality classes of services available from ISPs allow them to use the transmission capacity offered by ISPs more efficiently and thereby to offer more innovative products and services to end users. Content providers whose offerings depend to a high degree on the quality of delivery (e.g., cloud gaming, or real-time video broadcasting) could purchase a higher class of service and potentially ensure a high quality of experience for the end customers. On the other hand, content providers, whose offerings are not sensitive to occasional jitter or increased latency, could opt for a cheaper best effort delivery. This may lead to the emergence of new types of content classes, dependent on the high quality of services, which would not be available under a uniform quality regime of net neutrality regulation.

Of course, simple logic dictates that product differentiation is usually accompanied by price differences (see principle 3).[105] Price effects affect how the gains from the overall increase in welfare due to product differentiation are allocated across all market participants. Moreover, the choice to introduce or increase product differentiation can often be a strategic decision. In strategic settings, firms typically have incentives to differentiate their offerings from their rivals because it allows them to reduce competitive pressures and thus charge higher prices. But nevertheless from the perspective of the total welfare criterion, the introduction of product differentiation in non-strategic settings is unambiguously desirable. Because net neutrality regulation would in general have the effect of limiting (or in

---

104 See Spence (1976).

105 If two products of different qualities were offered at the same price, the product of inferior quality would be driven out of the market because all consumers would demand the superior good. That is why, if net neutrality regulation does not allow differential pricing, one does not expect service classes of differential qualities to arise.

its extreme from even totally prohibiting) product differentiation, it is likely to lower overall welfare.

"Content delivery networks" (CDNs), such as Akamai, provide quality of service enhancements to certain large content providers based on specialized technologies. Although, as we understand it, the data between a CDN and an end user is handled on a non-prioritized, "best effort" basis, the technology ultimately improves the end user's experience. Moreover, an increasing number of large content providers (such as Google) provide these services internally by creating a global network of local data centers ("server farms"). These facts indicate an existing demand for the (optional) services providing quality exceeding that of a best effort network, even in the environment in which ISPs cannot differentiate. Thus, quality differentiation appears important to end users. In this context, strict net neutrality regulation would in a sense discriminate against small and medium content providers who are too small to be (currently) served by CDN providers or to supply similar services in-house. Relaxing net regulation rules would allow the small and medium content providers to purchase higher quality of service, an option available right now through CDNs only to the largest CPs.[106]

Principle 3: Price discrimination increases total welfare

For an economist, price discrimination describes a practice of charging different buyers different **net prices** for the same product. In other words, price discrimination happens if differences in prices paid by buyers cannot be justified only by cost differences but also reflect differences in buyers' willingness or ability to pay. So the definition of price discrimination also describes the situation, in which two buyers pay the same price for a good or service even though the cost of serving them differs, for example, where all customers pay the same price regardless of their location.[107] In contrast, if the same good is sold at different prices to different consumers, but the price differences fully reflect the differences in costs (e.g., transportation costs), then for an economist there is no price discrimination.[108]

---

[106] Furthermore, net neutrality regulation might impact the CDN market in many different ways. In particular, net neutrality regulation in an extreme could also imply that CDNs have to be dissolved as they facilitate the prioritization of traffic. For more detail on this, see Ovum (2010).

[107] One can argue that this describes pricing on the Internet to a large extent today: prices are usually identical, but the cost of serving different customers can differ substantially.

[108] See Phlips (1983).

Differentiated products (see principle 2) can also be sold at discriminatory prices. If two varieties of a differentiated product are sold at prices which do not fully account for the differences in costs due to product differentiation, then price discrimination takes place. For a multiproduct firm, price discrimination often results from a joint profit maximization problem. Rather than optimizing the profits in each market segment separately, the firm takes into account several product models (or several time periods or regional markets) simultaneously and charges different (net) prices for each of them. For example, uniform delivery prices with transportation costs imply discrimination that favors more distant customers at the expense of customers located closer to the seller's plant.

Although price discrimination may invoke negative reactions and connotations among the public, it is a practice that is widespread in a variety of market settings. Price discrimination "might be as common in the marketplace as it is rare in the economic textbooks."[109] There is also a common understanding among the economic profession that it is generally welfare-enhancing. As a common practice in both concentrated and competitive settings, price discrimination only occasionally raises competition concerns and justifies per se prohibition.[110]

Price discrimination typically occurs in markets with differentiated products (see principle 2), but it is also possible in settings with homogeneous goods. In a market with perfectly competitive firms the law of one price prevails and price discrimination is impossible. Hence, it is often claimed that any firm that is able to engage in price discrimination must have market power, if only because some of the prices it charges are above the marginal costs. However, more recent research has shown that discriminatory pricing can also be prevalent in competitive industries and does not necessarily imply existence of market power. For example, it has been recently recognized that price discrimination may be a necessary feature in competitive industries in which there are high reoccurring fixed costs and barriers to entry are low.[111] If firms were to charge prices at marginal costs they would not be able to recover the high fixed costs and would end up consistently losing money. If they were to charge a high uniform price they would attract entry. So even sellers constrained by competitive conditions may find it necessary to engage in price discrimination as a way to recover fixed costs and

---

109 Phlips (1983), 7.

110 One notable exception is to input price discrimination between vertically integrated and independent downstream producers which may lead to margin squeeze and foreclosure.

111 See Baumol and Swanson (2003).

break even. Similar conclusions are reached in industries with large joint and common costs, which give rise to economies of scale and scope.[112]

Price discrimination is a common feature of more competitive, oligopolistic markets. The welfare implications of the monopolistic setting, however, do not necessarily apply. On the one hand, it is still the case that price discrimination tends to increase total industry output which increases efficiency. On the other hand, while monopolistic price discrimination benefited the producers and extracted consumer surplus, the opposite can be true in more competitive settings. Price discrimination in competitive settings can benefit consumers by intensifying competition among sellers. It is still generally true that – holding other things constant (including the behavior of its rivals) – an individual firm typically has incentives to discriminate.[113] However, if all firms were to switch from uniform to differentiated pricing, it is possible that overall profits in the industry would decrease. This often takes place as losses from intensified price competition due to price differentiation often exceed gains from additional surplus extraction allowed by price differentiation. The result is an increase in total and consumer welfare, but a reduction in firms' profits. If the firms could commit to uniform pricing, they would prefer to do so.

The output expansion effect of price discrimination (which is generally a feature of both monopolistic and oligopolistic settings) has an important additional implication in network industries and two-sided platforms such as the Internet (see also principle 4 and 6). In such settings, the output expansion due to price discrimination not only reduces the deadweight loss on one side of the market, but generates an additional positive effect because increased participation on one side is beneficial by increasing the value of the market or a platform to the other side. So the beneficial effects of price discrimination in two-sided markets are likely to be even greater than in standard one-sided markets.[114] The economic intuition behind this observation is very similar to the "waterbed effect" (discussed later as economic principle 4): If price discrimination increases participation on one side of the market, it generates a positive externality on the other side of the platform. Extending this intuition further, the greatest benefits can be achieved if price discrimination is possible not just on one, but on both sides of the market, as

---

112 See Levine (2002).

113 Ignoring potential strategic response from its rivals, a firm faces a standard optimization (profit maximization) problem. Allowing it to differentiate its pricing expands the set of feasible strategies which in turn might enable it to (possibly) reach a higher maximum.

114 See Weyl (2010).

increased participation on each side due to price discrimination increases demand on the other side and creates a virtuous feedback loop of enhanced benefits.

Price discrimination, through its output expansion effect, may also significantly impact incentives to invest and innovate. In a highly stylized setting, the return on innovation (such as a reduction in the marginal cost of production) is proportional to the level of output.[115] For example, consider a firm which sells 100 units at a mark-up of €10 at uniform price or 120 units at a mark-up of six with discriminatory pricing.[116] Suppose also that the firm may undertake an investment which would reduce its costs (and hence increase its mark-up) by one euro per unit. For any given cost of such an investment, it becomes more profitable and thus more likely, the more units the firm sells. Because we generally expect a higher level of output with price discrimination than with uniform pricing, incentives to invest and innovate should also be correspondingly higher with price discrimination than with uniform pricing. There are formal economic models which extend this highly stylized example. Taking the investment incentives explicitly into account they show that in some circumstances the welfare benefits associated with price discrimination may be sufficiently high so that even the consumers who are discriminated against can be better off. The intuition behind these results is that with price discrimination investment incentives are higher and so marginal costs of production are low, which may lead to a situation in which prices for both groups (in case of a third-degree price discrimination) of consumers are lower than a uniform price (with lower innovation and thus higher marginal cost) would be.[117]

To summarize, the overall effect of price discrimination on total welfare can be characterized as generally positive. The aggregate effect on consumers' welfare is less clear and some groups of consumers may be better off while others may be worse off. For example, consumers with a low valuation of the good usually are better off, since they are no longer priced out of the market, while consumers with a high valuation are typically worse off, since they may need to pay higher price. The more competitive the environment, however, the more likely consumers are better off.

---

[115] Profit is a product of quantity sold and (average) mark-up. Mark-up is the difference between the (average) price and cost.

[116] We are assuming here that, consistent with our earlier analysis of price discrimination in a competitive environment, mark-up and profitability is lower with price differentiation. However, the example would still be valid in the monopoly setting, where both mark-up and profits with price differentiation were higher because the calculation does not depend on the level of mark-up but rather on the increase in mark-up due to investment in cost reduction, which is the same regardless of the market structure.

[117] See Alexandrov and Deb (2010).

Principle 4: A price increase to content providers reduces the price to end users ("waterbed effect")

A well-established and quite general theoretical result in the literature on two-sided markets states that increasing prices for one side usually leads to lower prices for the other side.[118] This effect has important implications for net neutrality regulation, in terms of resource allocation within the industry. Abstracting for a moment from any potential changes to total welfare that may result from the ISP's ability to charge fees absent net neutrality regulation (we discuss these welfare effects in the following sections), such fees seem to imply a simple reallocation of resources from content providers towards the ISP. However, increases in fees for content providers also increases the margin an ISP can earn on content providers. This in turn implies that content providers become a more attractive market group for the ISP and it is going to try to attract this group. As content providers value the presence of end users on the other side, the ISP is going to try to attract more end users by lowering the prices to end users. Thus, increases in content provider fees are likely to generate incentives for ISPs to lower subscription fees for end users, as they try to capture some of the positive cross-market externalities on the other side of the market. Therefore, the allocative effect of the charges on content providers implies a (partial) transfer from content providers towards end users.[119]

---

[118] See for example Rochet and Tirole (2006).

[119] Irrespective of the level of competition between ISPs, ISPs have an incentive to price the side of the market who multi-homes (content provider) more than the side of the market that engages in single homing (end consumer). The intuition for this is that ISPs control access to end consumers for content providers. They therefore have an incentive to attract end consumers in order to exploit this control in relation to the content provider. Effectively, the content provider subsidizes the access of the end consumer to the platform, and prices for both market sides are biased in comparison to the social optimum. Some proponents of net neutrality have put this result forward in order to justify restrictions on the market price for content providers which might partially soften the bias. However, the theory to date is not conclusive on the effects of a restriction of prices to content providers. For example, the paper of Economides and Tag (2009): "Net Neutrality on the Internet: A Two-sided Market Analysis", Working Papers 07-27, New York University, Leonard N. Stern School of Business, Department of Economics, finds that termination charges decrease social welfare for particular parameter ranges of their model. This result, however, relies on a couple of essential assumptions: (1) consumer demand for Internet access is inelastic and (2) ISPs can only charge uniform termination charges to Internet content providers. Additionally, it is important to see that consumer welfare decreases with the prohibition of termination charges as this implies that consumer prices are going to increase and this decrease in utility due to higher prices is not compensated by the increase in content availability on the Internet. We propose to analyze those critical assumptions. For example, the first assumption appears to run

Note also that the waterbed effect does not allow determining the aggregate effect on both sides of the market, and theoretically overall revenues for ISPs from changed fees can be either higher, lower or unchanged. It is therefore impossible to determine purely on theoretical grounds the aggregated effect on ISPs from higher fees levied on content providers and lower fees for end users.

The ability of ISPs to extract fees from content providers can be severely limited by the bargaining power between the two types of agents. In fact, in the current environment, there are examples of content providers (e.g., American sport network ESPN) successfully demanding payments from ISPs for their content rather than the other way around. The bargaining power of ISPs is in particular determined by the attractiveness of its end user base. As the introduction of net neutrality regulation is unlikely to significantly impact the end user base, the distribution of bargaining power between the two sides is unlikely to be affected to an appreciable extent. Thus, in some cases the ability of ISPs to levy or increase fees on content providers may be very limited or non-existent.

Finally, if content providers have a direct (financial) relationship with end users (e.g., if end users pay for each video download), higher fees charged by ISPs to content providers can be, to some extent, circumvented and passed through to end users. In other words, the existence of a direct financial relationship between end users and content providers creates a possibility to circumvent or bypass the platform-pricing mechanism and thus may reduce the strength of the waterbed effect.

Principle 5: The difference in expected profitability with and without investment/innovation affects incentives to invest and innovate

The difference in expected profitability with and without investment or innovation affects incentives to invest and innovate. There are a number of important factors that can affect these expectations and hence influence future investments and innovation. Expected profitability depends to a large extent on the competitive environment. Uncontested monopolists have low incentives to invest and innovate in their core markets (so-called *fat-cat effect*).[120] If industry participants expect competitive conditions in the future and thus expect not to be able to appropriate

against the first-hand evidence where adoption rates of higher speed Internet access are particularly low in many European countries. Furthermore, the assumption of uniform termination charges appears to be unrealistic in a setting of bilateral negotiations. However, with differentiated termination charges it is significantly less likely that content is priced out of the Internet as ISPs have an incentive to keep content in order to be attractive for the end consumer (as well as for transit agreements).

120 See Fudenberg and Tirole (1984).

their innovation efforts, they will also have low incentives to invest and innovate, because they expect that profits from their innovation are going to be competed away.[121] Incentives to innovate are largest in highly contestable or oligopoly markets. Innovations allow firms to differentiate from each other and thus lessen competitive pressure or prevent rivals from "catching up." Strategic considerations may provide additional incentives to invest, for example, to deter entry or the expansion of rivals.

In some instances not only relative expected profitability with and without the investment can play a role in the investment decision, but also absolute profitability levels: if capital markets do not function perfectly, then companies might have to partly finance their investment out of their own profit.[122] Thus, companies with higher profitability might be the only ones investing. In the context of Internet provision, perfect capital markets might be affected by the following two issues. Firstly, asymmetric information might be an important factor for the capital markets. To the extent that ISPs have better information about the profitability of investing in capacity expansion than the capital markets (and which is difficult to communicate credibly to external investors), they might need to fund investments out of internal capital. Furthermore, ISPs have developed out of a regulated industry. This might be perceived as a risk in the sense that future regulatory intervention in an industry which is already on the radar screen of regulatory authorities is more likely than in other industries. This in turn implies that the expected profits of future investments might be capped from above decreasing in particular the expected profitability from the point of view of the capital markets.[123] Due to those features of the capital market, ISPs might find it hard to fund capacity expansion with external capital and might have to rely more on internal funding than, for example, content providers. These external funding difficulties might also be reflected in the development of share prices of prominent Internet service providers which have hardly increased since 2004.[124]

---

121 This provides the rationale for intellectual property rights (patents) awarded to innovators as those protect innovators from competition for a limited period of time sufficient to recoup the innovation costs. Patents thus increase incentives to invest.

122 For example, pharmaceutical markets exhibit a high and stable ratio of R&D expenditure to revenue of around 16%, see for example EU Industrial R&D Investment Scoreboard.

123 Of course, this risk also decreases the expected profitability from the point of view of the ISP. However, to the extent that the ISP has better information on the risk of regulation and if it can influence the risk to some degree, profitability is not decreased as much as from the point of view of investors.

124 See AT Kearney (2010).

Network effects (see principle 6), or more generally positive externalities (see principle 1), are often characterized as market failures which lead to suboptimal (too low) levels of investment and innovation.

Principle 6: Network industries benefit from interoperability

Network effects are somewhat similar to economies of scale: as the number of buyers and sellers both increase, the surplus available to each agent also increases. Therefore, the more members a network attracts the more value it generates for its members. Also, network effects often involve externalities, in the sense that prices do not fully incorporate the benefits of one person's entry into the network on existing members. This leads to the under-adoption of the network.[125]

Interoperability between different networks increases the size of the overall network available to end users and hence increases welfare. For example, compatibility in the form of interconnection – so that a phone-call originating on one network can be completed on another – is a fundamental principle of telecommunication systems and increases welfare by allowing more people to communicate with each other without a need to participate in multiple networks. Interoperability can be achieved through individual agreements between network operators or – if the transaction costs of such agreements are prohibitively high – more efficiently through standardization.

Standards can be formally agreed upon in an open standard setting process involving all interested market participants. Another possibility, especially likely in dynamic industries, is "bandwagon" or *de facto* standardization, where the standard adopted by the incumbent operator and/or early movers' determine later adoptions by end users (and competitors).

If consumer preferences are sufficiently heterogeneous, standardization may have welfare-reducing effects and thus involves a trade-off. If each group adopts its preferred standard, then each group enjoys the benefit of using its preferred standard but foregoes the benefit of being part of a larger network. On the other hand, if two groups of consumers prefer different standards and if a single standard is adopted, then both groups enjoy the network benefit of being part of a larger network, but one of the groups foregoes the benefit of using its preferred standard. In other words, one of the primary costs of standardization is loss in product variety (see principle 2). Whether a single standard is optimal or not

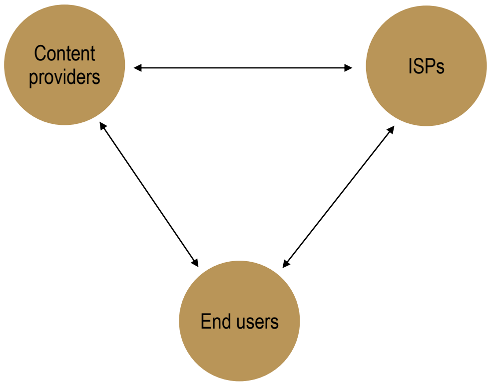
[125] See Klemperer and Farrell (2008).

depends on which one of the two effects dominates, that is, whether the network effect or product differentiation effect on utility is larger.[126]

Platform operators' incentives towards interoperability depend on their size. If firms can prevent the adoption of a standard by their rivals, multiple standards may occur due to coordination breakdowns, even though they would all prefer to coordinate: Large platform operators usually have incentives against interoperability, because competing networks are small, so their users gain little from interoperability while their rivals' users gain a lot. By preventing interoperability, they may induce users of the competing networks to join the dominant network. In contrast, small platform operators usually have strong incentives for interoperability, because the ability to connect with the large network generates a large benefit for their members.

In the context of the Internet, at least three different types of externalities accrue to all participating agents (see figure 6).

Figure 6: Externalities between Internet participants



Source: ESMT CA.

When ISPs invest in their infrastructure or content providers develop new content they often generate surplus that neither the ISPs nor content providers fully capture through their fees, creating an externality. For example, increased bandwidth capacity may stimulate development of the innovative type of content not feasible at lower capacities. It may also attract new users or encourage already connected users to consume more. Innovative content may stimulate demand for broadband Internet infrastructure capacity and access and it may also attract more end users to the Internet. Finally, increased participation by end users provides additional incentives for content creators and ISPs to expand their offerings.

[126] See for example Yoo (2008).

All three types of positive spillovers feed on each other and lead to the creation of a positive feedback loop. A recent FCC document describes the cumulative effect in the following way:

> Networks, devices and applications drive each other in a virtuous cycle. If networks are fast, reliable and widely available, companies produce more powerful, more capable devices to connect to those networks. These devices, in turn, encourage innovators and entrepreneurs to develop exciting applications and content. These new applications draw interest among end-users, bring new users online and increase use among those who already subscribe to broadband services. This growth in the broadband ecosystem reinforces the cycle, encouraging service providers to boost the speed, functionality and reach of their networks.
>
> FCC (2010). Connecting America: The national broadband plan.

Beside interoperability, offering reduced or differentiated subscription prices to broadband end users (see principles 2 and 4) would also increase participation rate and thus strengthen the positive network effect.

Principle 7: Economic decisions involve trade-offs

Economic decisions usually involve making a trade-off. The first of ten commandments in the undergraduate microeconomics textbook by Mankiw and Taylor (2010) reads “people face trade offs.”[127] This also applies to regulatory decisions which affect how business is carried out on the Internet. Overall, the main opportunities associated with a limited enforcement of net neutrality are seen in better infrastructure utilization, higher incentives to invest into future generations of infrastructure as well as increases in consumer welfare as consumer prices decrease. The major economic risks identified in the debate relate to the fragmentation of content on the Internet, the ability of incumbents to foreclose competing content providers via price discrimination, and reduced innovation in services and content due to the redistribution of revenue streams to ISPs. In the following we will discuss some of the main trade-offs.

Trade-off 1: Consumer benefits from lower prices today versus consumer benefits from new content-related products and services tomorrow

In two-sided markets platforms balance their pricing between the two groups of customers. Simply speaking, increases in prices paid by content providers result in decreases in prices for consumers. This is reflected in principle 4 (waterbed effect). Thus, any business model impacting the price for content providers will

---

[127] See Mankiw and Taylor (2010), 4.

eventually trigger price changes for consumers. This simple mechanism explains the first trade-off as those business models which imply increases in content providers' prices might result in lower prices for consumers today but at the same time might reduce the incentive to innovate for content providers. This will eventually lead to less variety of products tomorrow.[128] This trade-off is nicely summarized by the quote from Lee und Wu (2009):

> It is an open question whether, in subsidizing content, the welfare gains from the invention of the next killer app or the addition of new content offset the price reductions consumers might otherwise enjoy or the benefit of expanding service to new users.[129]

Trade-off 2: High quality of service for some versus average quality of service for all content providers or end users

Allowing traffic management to install priority lanes for time-sensitive traffic (or more generally increase the quality of service for some) will benefit those applications or end users that can take advantage of those priority lanes and might hurt others due to the potential decrease of the average speed outside priority lanes.[130]

Trade-off 3: Incentives to innovate in content and services versus incentives to invest/innovate in infrastructure provision (for non-complementary network and content investments)

Investment incentives for ISPs are usually discussed in relation to the effects on innovation incentives for content providers resulting from a change in the pricing scheme. The simple intuition is that for the ISP the incentive to invest depends on

---

[128] It should be noted that some of the future products and services might be very data intense (e.g., video applications). Therefore, this trade-off might be limited in the sense that some applications which might only be realized if no additional charges are levied on content providers might not meet the required infrastructure capacity if ISPs have no sufficient incentives to invest due to low content provider prices. This effect is illustrated in trade-off 3.

[129] Lee und Wu (2009), 67.

[130] This trade-off is depicted in the static analysis of the paper by Krämer and Wiewiorra (2010). They analyze the effects of a move from uniform to differentiated prices for different types of time-sensitive applications/customers. They find that in the short run (with given infrastructure) the efficiency of the bandwidth allocation increases as customers with a higher utility for fast delivery have lower delay rates whereas the delay for less sensitive customers increases. Thus, in this model the less sensitive customers are worse off and the more sensitive customers are better off with differentiated prices. A judgment on the desirability of such changes depends on how those benefits and costs are aggregated.

how much profit it can generate with the investment. The profit in turn depends on how much it can charge its users and in particular content providers. At the same time, for content providers the incentive to innovate also depends on how much profit they are able to make with the innovation. If content providers are charged higher prices, it is less attractive to innovate. Thus, there appears to be a trade-off between incentives to invest for ISPs and incentives to innovate for content providers.

The above argument focuses on the transfer of rents between ISPs and content providers and the resulting profit expectations and investment trade-off. However, it is also conceivable that there is a certain degree of complementary between those two types of investment decisions as certain applications might be facilitated by investment and innovation in the infrastructure. Currently, during congestion periods quality-sensitive applications might not function and thus be crowded out of the market. Higher incentives for ISPs to invest in infrastructure expansion in combination with higher quality offers of ISPs might open the market for content providers and thus stimulate investment, thereby effectively resolving the trade-off.

The focus of the emerging literature has not been on the complementarity between ISP investment and innovation in content provision. However, there are some papers which (indirectly) model such a positive link by assuming that content providers are more profitable on a higher quality platform, for example, because quality of transmission increases advertising revenues.[131]

Trade-off 4: Net benefits of ex ante versus ex post regulation (antitrust enforcement)

Finally, another important trade-off implies the cost of ex ante regulation versus ex post regulation. This addresses the more general question of whether the risks of the anti-competitive effects of business models (e.g., free pricing) are sufficiently high to justify a regulatory intervention. On the evaluation of this trade-off, the UK regulator Ofcom recently argued that:

> Generally speaking, our initial position is that discriminatory behaviour is only a potential issue where firms have substantial 'market power' and could discriminate in favour of their own services. In this case, any form of discrimination will come under very close scrutiny to ensure that there are no anti-competitive effects. We believe that there is insufficient

[131] See Krämer and Wiewiorra (2010), or Njoroge et al (2010).

evidence at present to justify the setting of blanket restrictions on all forms of traffic management.[132]

4.2. Detailed analysis of proposed business models

This section analyzes the proposed business models along the lines of the general economic first principles taking into account particular aspects of the Internet market. The analysis considers the effects of different business models in comparison to the outcome resulting from the current business model as described in section 2.2.

4.2.1. BM 0: Congestion-Based Model

In the “Congestion-Based Model” the network still operates under the current best effort principle, but users on both ends of the network are (partially) charged for congestion externalities they create. For end users, this means that they can choose between different flat rates with differentiated data caps. For directly connected content providers and transit ISP partners, an element of peak-load pricing is introduced. Peering agreements would, however, not be altered.

In particular due to the general interoperability, currently the Internet exhibits properties of a **common pool** where content and users cannot be excluded. This in turn results in an overuse of the common-pool resource. The literature on common pools discusses a variety of options available to solve the over-usage problem, among them introducing appropriate pricing structures. In this respect congestion-type pricing has two effects:

- **End user volume pricing:** Flat rates create incentives for all to consume more and hence in consequence uniform flat rates encourage congestion. Volume caps limit the total amount of traffic generated by end users.[133] Furthermore, differentiated volume caps create positive long-term marginal prices and thus reduce incentives for excessive consumption present currently at zero marginal price. This addresses the overall congestion issue, but has no effect on adjusting allocation of traffic over the different periods of the day.
- **Peak-load transit pricing:** For a given level of infrastructure, more congestion-based pricing softens congestion problems to the extent that

---

132 See Ofcom (2010). Traffic Management and ‘net neutrality’ – A Discussion Document, 2.

133 In the classical common-pool example of fisheries this measure would amount to allocating (different) quotas to each fisher.

consumption can be smoothed over time. If congestion only arises during a limited period of the day, then creating incentives to reduce consumption during those peak periods might be sufficient to avoid congestion. As electricity and other markets demonstrate, peak-load pricing can be an efficient tool to allocate traffic from peak periods to non-peak periods and utilize the existing infrastructure more efficiently.

Thus, introducing more volume and congestion-based pricing ameliorates congestion externalities arising due to the common-pool characteristics of the Internet. However, a number of issues still persist:

- **Heterogeneity of content:** Peak-load pricing could be an efficient tool to eliminate congestion externalities if content was homogenous not only at the data packet level, but also on the service level. However, content differs with respect to data rate, quality sensitivity and also the economic value. An "average" peak-load pricing system might ameliorate congestion in peak times. However, the effect might be too small for highly quality-sensitive content. Thus, without incorporating in particular differences according to the quality sensitivity of content, some sensitive but valuable services might be crowded-out by quality-insensitive services which might have lower economic value. Also, on the lower end of quality sensitivity, some content might inefficiently exit the network due to the fact that it cannot pay the peak-load prices. In this sense, peak-load pricing is unlikely to eliminate the congestion problem.[134]
- **Persistence of peering agreements:** The business model postulates that peering agreements persist. Thus, for peering traffic common-pool problems remain as peering traffic cannot be effectively excluded. As tier 1 peering partners generate a significant amount of traffic, the persistence of peering agreements puts a serious limitation to the effectiveness of the new pricing system to solve congestion issues.
- **Ability to steer traffic by content providers:** With the (partial) introduction of peak-load pricing, directly connected content providers as well as indirectly connected content providers who experience peak-load pricing through a system of cascading payments have an incentive to avoid peak traffic. The question is to what extent they are actually able to do so – given that most of the traffic generation originates with the end user. We see a number of ways of how the content provider could influence the timing of traffic generation: content providers could try to

134 See for example Kruse (2009).

influence consumer behavior directly either by reducing the availability of their services during peak hours or by offering different pricing options for individual transactions based on the time of the day.[135] Content providers could also reduce peak traffic by increasing their demand for CDN services or by designing less traffic-intense applications. Those services reduce overall traffic and might thus also help in reducing traffic during peak periods. Overall, peak-load pricing for content providers is only going to soften congestion externalities. However, since content providers are unlikely to be able to perfectly steer end users' behavior, it is unlikely to entirely internalize the externality.

In summary, peak-load pricing for content providers and volume caps for end users may reduce the congestion problems by smoothing the network consumption over different times of a day and increasing the average utilization of the existing capacities, but it would not eliminate congestion entirely.

The effect on (average) prices for different type of content providers is not obvious. For example, a peak-load pricing system could be designed in such a way that the average price for all content providers stays the same. This would imply that content providers whose content is mainly accessed outside the peak periods would enjoy lower prices whereas content providers those content is mainly accessed during peak periods suffer from higher prices. Consequently, some content providers which are unable to steer their traffic effectively and those whose economic value is small might exit the market. At the same time, content providers which were previously kept out of the market by high congestion can now enter the market of reduced congestion with delay-sensitive content.

With respect to the **incentives to invest,** more efficient usage of existing infrastructure stimulates ceteris paribus more efficient investment incentives. However, this does not necessarily mean that absolute investments increase in comparison to the status quo. If more complex pricing results in the reduction of peak traffic, then overall less capacity may be needed. This in turn may imply that the necessary investment decreases in comparison to the scenario where peak

[135] For example, a video download service might not be available from 7:00 p.m. to 11:00 p.m. According to Deutsche Telekom, this type of traffic steering is, for example, applied on Google's infrastructure in the sense that HD YouTube videos are not accessible during peak hours. Alternatively, the same video offer could come at a discounted price if the end user accepts the delivery of the download within the next 24 hours. This arrangement would give the content provider the flexibility to initiate the traffic at the most cost-effective time.

traffic is not reduced. Consequently, overall traffic requirements can be satisfied at lower investment levels.[136]

In terms of **price differentiation,** different volume caps associated with different flat rates display an element of price differentiation in that average prices for end users can differ. However, already to date, effective average prices differ in the following sense: suppose a single flat rate is offered to all end users, some of which are very heavy users whereas others only occasionally use online services from home. Because all users pay the same flat fee, the heavy users in effect pay a lower average price than the light users, so in a sense light users subsidize heavy users. In contrast, the introduction of different volume caps offered at different “flat rates” can mitigate this cross-subsidy as light users are likely to switch to the cheaper flat rates with lower volume caps. The menu of flat rates offered could, in principle, be designed so that the expected average price per bit for light users equals the expected average price per bit for heavy users, reducing the cross-subsidy between users.

To understand the likely effects of price differentiation it is useful to start with a simplifying scenario; that is one in which a monopolistic ISP could identify light and heavy users and “force” them into certain packages of volumes and flat rates (i.e., first-degree price discrimination). That way light users, who are supposedly more price-sensitive, would likely benefit from lower prices compared to the single flat rate. Indeed, lower prices are also likely to attract new light users who were previously not connected to broadband services. It should be emphasized that high prices are the main constraint to high broadband penetration in many countries; for example, Home Broadband Adoption 2009 found that 32 percent of American dial-up users said that broadband prices would have to fall for them to switch to broadband.[137] On the other hand, price differentiation would likely increase average prices for heavy users because now ISPs price toward their willingness to pay. Hence, ISPs charge both light and heavy users according to their willingness to pay, that is, relatively low or high averages prices, respectively.

However, in practice ISPs will not be able to identify and force light and heavy users into certain packages of volumes and flat rates. Light and heavy users can

---

136 The smoothing effect of peak pricing means that peak traffic is reduced, but there may also be a change in the overall level of usage. If overall usage increases more than peak traffic is smoothed out then there may still be increased peak traffic and more investment needed.

137 Another 20% said nothing would get them to switch, 17% cited availability as the obstacle to switching, 16% stated they did not know what would get them to switch and 13% cited some other reason.

simply pick a tariff scheme that suits them best; that is, they self-select a certain plan. In such a situation, a profit-maximizing monopolistic ISP faces a number of constraints. First, because each of the users can choose any plan they want, for the screening to be successful the ISP needs to ensure that the heavy usage plan is preferable to a heavy user and a low usage plan is preferable to a light user. Additionally, both prices need to also be sufficiently low to encourage the participation of both types of users. Standard economic theory suggests that the entire rent of the low-value user is extracted or otherwise the monopolist could simply increase the low-tariff price and increase its profit. Furthermore, the heavy user must be (almost) indifferent between choosing both plans or otherwise the monopolist could increase its profit by increasing the price of the high value plan. This, however, implies that some surplus is left to the high-value user type because otherwise he would switch to the low value plan. In an environment where user groups differ substantially, the above screening mechanism might not be relevant. Given the widely differing consumption patterns on the Internet (see also fact 4 in section 2.3), it can be assumed that heavy users are unlikely to ever want to accept any offer that is made to the light user.

The effect of price differentiation on users is ambiguous. Light users may benefit from lower total price and the most price-sensitive users, currently priced out of the market, may benefit by being able to subscribe to broadband at the lowest package price. Heavy users may be worse off and end-up paying more than under single flat rates. However, the overall welfare implications of this type of volume discounts are considered to be positive as they tend to increase participation and output.

Within the best effort network (technical) **interoperability** is entirely preserved. However, the pricing mechanism for content providers introduces another set of issues related to interoperability, in particular if different ISPs apply different schemes of peak-load pricing. Consider an example where one ISP prices according to peak hours during the day and another ISP who only differentiates its pricing according to weekdays and weekends. In this situation, in order to steer the end user demand, the content provider would ideally change its content availability and pricing strategy according to whether end users are connected over one or the other ISP. This may significantly increase the complexity of end user pricing and furthermore might also be considered as a violation of neutrality principles.

Furthermore, content is likely to travel through different networks sequentially. Thus, if different networks apply different peak-load pricing mechanisms, this might significantly increase the complexity of how to steer traffic for the content provider.

To the extent that the two different peak-load pricing systems reflect the different patterns of congestion, it might be useful to implement different systems. However, at the same time those different systems imply higher costs as the steering of end user traffic becomes more complex.[138] If these transaction costs are prohibitively high, then introducing peak-load pricing may not be an attractive practical option.

Finally – unrelated to the discussion of the economic first principles within the proposed business model – a vertically integrated player might have an incentive to engage in foreclosure. This **risk of foreclosure** is at the core of the net neutrality debate in the US. The fear of proponents of net neutrality is that absent some specific rules, vertically-integrated ISPs might discriminate against certain content providers which are in competition with services of the ISP. Generally, competition law can and does deal with foreclosure on an ex post basis. In order to justify regulation on the grounds of foreclosure concerns, one needs to specify why a particular industry is particularly prone to foreclosure and that an ex ante regulation would thus be likely to have net benefits. In particular, foreclosure concerns depend on the level of concentration of the industry. Since Europe has implemented access regulation, the level of competition for end users is presumably larger than in the US. Furthermore, it appears that in the US American ISPs are vertically integrated to a significantly larger degree than their European counterparts, in particular in data-intense video applications. This partly results from the fact that traditional TV cable operators have moved into the provision of Internet services and thus have a strong grip on video content.

In comparison to the US market, the European ISP market is less concentrated and vertically integrated to a lesser extent. Consequently, foreclosure concerns are less relevant in the European debate. Nevertheless, in Europe, too, some concerns might remain as access regulation might also miss the target level of competition. With significant market power foreclosure concerns might actually be grave in the industry as the content market in particular also displays strong elements of network effects. Furthermore, the detection of anti-competitive discrimination on the technical level might be very hard to detect.[139]

With respect to such foreclosure concerns, the “Congestion-Based Model” does not change the situation significantly compared to the current status quo. The ISP might still implement some reasonable network management techniques in order to

---

138 Furthermore, standardization of pricing schemes between different eyeball ISPs might bear some risks of collusive behavior.

139 See for example the case of Comcast against BitTorrent.

handle traffic during peak hours as in the status quo. The difference might be that due to congestion-based pricing, congestion is eased and the need for network management reduced. Still, network management techniques are likely to remain a part of the toolbox of running the network and could be misused by a dominant vertically integrated ISP to foreclose content competitors. However, it should be emphasized that so far the evidence on potential foreclosure cases using network management techniques, even in the US market, is limited. There are two prominent cases.[140] Comcast, a large cable ISP in the US, allegedly secretly discriminated peer-to-peer (P2P) applications such as BitTorrent. The other case concerns Madison River, a regional broadband provider, who blocked VoIP applications such as Vonage for its DLS subscribers. Once the practice became public, both companies changed its network management strategies even though it is uncertain whether the FCC has the right to demand modifications. This evidence points to the importance of transparency in the market but does not provide strong support for a generally high risk of foreclosure in the industry.[141]

Furthermore, the additional pricing element where prices are differentiated with respect to different periods of the day might supply ISPs with another means to discriminate competing content providers. However, it might be argued that contracting terms might be easier to compare and discriminatory behavior might thus be easier to detect.[142]

As in the status quo, one major requirement for functioning competition in the provision of Internet services to be effective relates to the level of transparency over the use of network management tools for end users and content providers.

---

[140] See FCC (2009).

[141] Foreclosure concerns related to the ability to steer or "manage" traffic cannot be totally dismissed as evidenced by the recent opening of an investigation by the European Commission into Google's search algorithm and online practices. The investigation was initiated by complaints from Google's competitors who alleged that Google maintains a dominant position in online search services and manipulates its search results to favor its own services over services by competing providers. This type of behavior violates "search neutrality." By analogy, similar behavior (i.e., prioritization) of traffic by a dominant ISP provider favoring its own vertically-integrated (or exclusive) online content over competing providers' content would also attract scrutiny from competition authorities. To date, the crucial difference between the recent Google case and ISPs is that the ISP market is subject to access regulation and therewith to a significant degree of competition (as long as access regulation is properly functioning).

[142] Additionally, it should be mentioned that it is not unlikely that future industry regulation incorporates the notion of non-discriminatory pricing (see for example BMWI, 2010). This would likely incorporate the presumption that discriminatory pricing causes consumer harm. This context facilitates enforcement of non-discrimination and makes foreclosure more difficult for dominant vertically-integrated players.

Table 11 summarizes the discussion of effects. The take-away points are that the "Congestion-Based Model" reduces congestion and allows more efficient utilization of the existing infrastructure in comparison to the current status quo. However, it is unlikely to provide sufficient incentives to eliminate congestion entirely. A minor drawback is that its implementation can lead to increased complexity for content providers as well as subsequently end users.

Table 10: Impact assessment of Congestion-Based Model (BM 0)

		Possible impact assessment
Congestion		Likely decrease in congestion during peak periods Some smoothing of bandwidth consumption over time
End users	**Prices**	Ambiguous: Heavy users may pay more (according to higher volume) Light users may benefit from lower total price
	Participation	Increase in participation due to currently non-participating light users who may subscribe at lower prices and due to improvement in congestion which might induce modest increase too
Content providers	**Prices**	Ambiguous: Likely increase in prices for CP unable to shape traffic Prices for other CPs unchanged or possibly decreased
	Participation /Innovation	Ambiguous: Less crowding-out of delay-sensitive content Potentially some exit by content providers which cannot steer traffic Increased incentives to invest in technologies to shape traffic. Possible decrease in incentives to invest in innovative content that cannot be shaped
ISPs' incentives to invest		Capacity is utilized more efficiently which induces more efficient investment incentives (Not necessarily more absolute investment)
Regulatory costs	**Ex ante intervention**	Transparency for content providers and end users wrt network management tools (as in status quo)
	Ex post intervention (competition concerns)	No significant increase in the risk of foreclosure by dominant vertically integrated ISPs
Interoperability		Limited coordination problems related to different peak-load schemes of various interacting ISP

Source: ESMT CA.

4.2.2. BM 1: Best Effort Plus

“Best Effort Plus” preserves the best effort Internet, but introduces additional “priority lanes” for clearly specified, innovative services. While nothing changes in the pricing of the best effort network, innovative content providers negotiate bilaterally with the eyeball ISP and receive thus individualized price levels. Those bilateral negotiations might also imply a higher degree of vertical integration of ISPs into content provision. In contrast to the previous business model, the “Best Effort Plus” does not tackle congestion issues explicitly. Instead, the central focus is on the infrastructure necessary to facilitate the new services. Thus, it deals with the congestion issue indirectly in the sense that it assumes that certain online services cannot be provided over the existing infrastructure due to congestion. The central concerns relate to the impact of the innovative services network on the best effort as well as the risk for foreclosure in the presence of vertically integrated ISPs.

The **common-pool** perspective on the Internet provides an explanation as to why congestion issues arise in the best effort Internet. In contrast, innovative services do not exhibit the characteristics of a common pool: there is a clear contractual relationship between an innovative content provider and the ISP on the required quality of transmission. This quality guarantee is either only relevant on the network of the eyeball ISP or alternatively on a number of participating ISPs. However, there is no general interoperability agreement between different ISPs on the delivery quality of innovative services. Thus, there is no general free riding incentive as clear appropriation rules are set for the innovative services. However, both types of content – traditional and innovative – ultimately run on the same physical infrastructure - even though additional capacities might be deployed especially for new services. Thus, to the extent that part of the infrastructure capacity is allocated to innovative services, the congestion problem in the traditional best effort services might be aggravated.

The question is thus whether one expects a reduction in the transmission quality and an increase in perceived congestion in best effort after introducing innovative services on the same infrastructure. Ignoring investment incentives, if new content is put online which is accessed during peak hours this necessarily implies that the current content is experiencing stronger congestion externalities. If on top of this the innovative service receives priority access, then the quality of transmission on the best effort network is further reduced. The priority lane thus exerts an extra congestion cost on top of the standard cost which comes with the introduction of a new service.

Suppose it would be possible to define the optimal growth path for best effort on the basis of traffic growth and technological developments. In this case, it would also be possible to require ISPs to ensure that this optimal growth of the best effort network is achieved (dynamic best effort definition). However, this is difficult because the forecast of traffic as well as technological possibilities and associated costs of network provision are highly uncertain. To the extent that the optimal best effort network cannot be forecasted, it cannot be ruled out that innovative services reduce the transmission quality in best effort networks:

- **Content provider's choice of quality level:** The literature on net neutrality discusses the incentive of ISPs with market power to degrade the quality in the best effort network, for example, by deliberately routing traffic inefficiently and thus increasing delay in order to increase the attractiveness of the higher quality classes.[143] In the "Best Effort Plus" business model, content providers cannot freely choose the quality regime. Instead traditional services have to use the best effort and only innovative services – defined among other things by the quality requirements of the service itself – get access to priority lanes. Thus, even if ISPs were to strategically degrade the quality of the best effort network, they would not profit because the content providers would not be able to freely upgrade to the paid priority lanes.[144] In fact, due to the inaccessibility of higher quality classes for traditional content, the demand for best effort quality is foremost independent of the quality of transmission in innovative services. In this sense the markets for traditional and innovative services are independent. Consequently, the price and quality level is going to be independent just as the incentive to invest in the infrastructure for both types of traffic.
- **Degree of substitution between traditional and innovative content providers:** However, if innovative and traditional services are sufficiently close substitutes, then an enhanced quality of experience for the competing innovative service is going to reduce the attractiveness of the traditional service and its profitability. An ISP with significant market

---

143 See for example Choi and Kim (2010), Cheng et al (2010) or Hermalin and Katz (2007). This literature deals with the quality choice of ISP if quality differentiation is allowed and compares the level of investment into infrastructure with and without quality differentiation. It thus does not deal with foreclosure or "active degradation" incentives for dominant vertically integrated ISPs.

144 Instead of degrading the best effort quality network, the ISP might have strong incentives to lobby for a regulatory definition of innovative services which is broad and includes a number of new services.

power might take this externality of innovative services on traditional services into account when investing into the infrastructure. This bias would depend on the intensity of competition between traditional and innovative content providers: the more intense, the more profitable it might be for the ISP to bias the investment decisions. However, the nature of innovative services is such that they present a new, innovative service. Thus, by definition the service does not or only mildly compete with other services. It could compete, for example, on the level of attention of end users with other services. In this sense, there might be competition, however, if innovative services are appropriately defined, there does not seem to be a direct link.

- **Intensity of ISP competition:** Only ISPs with market power could eventually experience an incentive to crowd-out investment for best effort in favor of investment for innovative services. If the infrastructure provision for best effort is competitive, crowding-out effects are not to be expected.

In fact, the infrastructure deployed specifically for innovative services might exhibit positive externalities on the best effort level. This could be the case if usage patterns for traditional and innovative services are significantly different. For example, an innovative service might be directed at business clients (e.g., 3D video conferencing) and thus generate peak traffic during daytime. Traditional services might then use the available installed capacities during off-peak periods of innovative services which might coincide with the peak period of traditional end user services. Another channel for positive externalities could be found in quality guarantees for innovative content which also covers extreme traffic peaks. Because installed capacity would need to guarantee sufficient quality in extreme traffic peaks and since extreme traffic peaks only occur seldom, the rest of the time the free capacity could be used for traditional services.

In summary, we expect to observe limited crowding out of investment for best effort. Instead, extensive provision of infrastructure for innovative services might have a positive externality for best effort in the sense that the best effort traffic can use this extra infrastructure in case it is not busy with innovative services. The above discussion however assumes that the definition of traditional and innovative services is clear cut and only captures some truly innovative services which would not exist absent access to high-quality transmission. As soon as there is a degree of freedom in choosing to become an innovative service and/or the definition does not capture the innovative nature of services, the concerns of crowding out and the deliberate degradation of best effort quality as discussed in the economic literature might gain importance.

Generally, ISPs can offer more **differentiated qualities** to end users and content providers. Quality differentiation tends to increase consumer welfare (see principle 2 in the previous section). In this case, this is in particular so as the innovative service of the content provider would, by definition, not exist in the market if it did not enjoy the differentiated quality of transmission offered by the ISP. Thus, quality differentiation enables new, innovative services for end users.

Clearly, higher quality levels are going to be marketed at different prices to the best effort transmission. In the "Best Effort Plus" scenario it is assumed that individual bargaining occurs between the ISP and the content provider. Generally, commercial negotiations are regarded as an efficient pricing mechanism.[145] This has a number of effects. First of all, economic theory predicts that in many cases bargaining will be efficient. This would imply that all socially desirable innovative services would be facilitated. From this perspective, different prices for different quality levels will have no impact on the overall efficiency, but may affect the distribution of available rents between ISPs and content providers.

Bilateral negotiations are likely to result in individualized pricing structures. Anything could arise, from a two-part tariff specifying a fixed access fee and a traffic-based variable fee to revenue sharing agreement where bargaining partners share the joint surplus according to a percentage rule. Who of the two bargaining partners is going to receive the larger share of the surplus depends on who has greater bargaining power. The bargaining power is affected by many factors. For example, which party has the better outside options to the trade or which party holds superior information. In the present context, bargaining power can be affected, for instance, by the ability to implement a billing system more efficiently. The existing commercial relationships with end users through the provision of best effort Internet services might give the ISP a competitive edge when it comes to the marketing of the innovative services to end users who may choose their ISP as a trustworthy online payment partner.

As a consequence of bilateral bargaining some innovative services might have exclusivity arrangements with ISPs. There is a large economics literature dealing with exclusivity issues. In a nutshell, the antitrust concern associated with exclusivity agreements is that a potential competitor is kept out of the market. Generally, the parties have no incentive to agree to an exclusive contract if this implies that an efficient competitor of the input is excluded from the market. This

[145] The OECD, for example, acknowledges: "*Commercial negotiations have been shown to be the best way to deal with the many thousands of agreements which need to be transacted in order for the Internet to efficiently route traffic at the lowest unit cost for all parties.*" See OECD (2006), 9.

Chicago Critique fails when externalities are present: for example, one side of the market is fragmented and cannot coordinate[146] or the two market sides can coordinate with exclusive dealing in order to expropriate the more efficient entrant.[147] This implies that exclusionary contracts might be harmful to consumers in some special circumstances. This, however, also hints at the fact that a general prohibition of exclusive contracts might be over-inclusive and prohibits cases where this type of vertical agreement induces efficiencies (see further below).

In the present context, exclusivity could relate to two different layers of the supply chain: the ISP can either exclude other content providers or the content providers can exclude other ISPs.[148] The first type of exclusivity seems to be of limited interest if innovative services are appropriately defined. First, the definition of an innovative service implies that at least at the beginning little competition with other innovative services exist. Second, exclusivity on one network is unlikely to be sufficient to prevent a competitor from entering on a different ISP network. But striking exclusivity agreements with a critical mass of ISPs appears difficult. Thus, in summary it appears unlikely that exclusive agreements would lead to entry deterrence of competitors of innovative services. However, it cannot be ruled out and, as already pointed out, depends on the appropriate definition of innovative services.

Another concern of exclusivity agreements in relation to the innovative process is that it might be easier for larger content provider than for smaller content provider to strike an agreement with an ISP. This might for example be the case if larger content providers have lower administrative costs of striking exclusivity agreements. This might be seen as particularly alarming if one takes the stance that drastic innovations are more likely to be generated by smaller content providers than by larger ones. However, if there is a race of innovation, then the core question would focus on the probability that both content providers – small and large – find the innovation at the same time. This event might have a rather

---

[146] See Rasmusen et al (1991) and Segal and Whinston (2000).

[147] See Aghion and Bolton (1987).

[148] One can think of two interpretations: A content provider and an ISP could simply agree that no other content provider offering similar innovative services in the future is allowed onto the network of the ISP. An alternative way would be to think of an exclusive level of quality agreed upon between the ISP and the innovative service provider. To the extent that an innovative service depends upon a specific quality level, this type of agreement amounts to an exclusion of similar types of services on the same network. However, there appear to be many different technical ways to achieve a similar level of quality transmission. It is thus difficult to construct an exclusive quality class. A competitor might adapt its technical needs such that it can still enter the market not using the exclusive quality class.

small probability and the effect on the expected profitability might be minor. This potential effect of exclusivity agreements would have to be weighted in particular against the reduced incentives to invest in the absence of higher quality offers as in the status quo.[149]

The above argument looks at first-generation type of innovations, that is, first-time innovation. Another question focuses on the implications of the outcome of the first stage of innovation on subsequent innovative behavior: In a situation where exclusive agreements prevail, it follows that for innovative services only a limited number of content providers are in the market. This limitation of players on the higher quality infrastructure could potentially hamper the innovation process if successful innovation requires the ability to experiment on the higher quality infrastructure. However, the question arises why an ISP would have an incentive to write exclusive agreements in such a scenario: generally, competitive ISPs have an incentive to stimulate the innovative process in order to attract end users to their platform. If the innovative process is hampered when only a very limited amount of content providers have access to the higher quality infrastructure, then the ISP has ceteris paribus no incentive to agree to exclusivity. Thus, this discussion leads back to the question of foreclosure by a dominant vertically integrated ISP and again to the question of successful access regulation for ISPs.

Now, we turn to the question of whether exclusivity of content is likely to have anti-competitive effects regarding other ISPs. The question can be split into two sub-questions: (1) is competition/entry in the infrastructure provision for innovative services potentially hindered or prevented and (2) is the same true for competition and entry in the best effort network?

---

[149] The following argument draws from insights of the literature on patent races, see for example Tirole (1988): Suppose that smaller content providers are more likely to innovate drastically and assume that there is a large and a small content provider which are investing in R&D. The first company who successfully innovates would get the exclusivity agreement with the ISP. This in turn would imply that the investment efforts of the second company were futile. First of all, the assumption that a drastic innovation is more likely discovered by the smaller content provider would imply that the smaller content provider is more likely to be the first to innovate and strike the exclusivity agreement. However, in a situation where both content providers discover an innovation at the same time, the concern would be that the larger content provider would strike the agreement with the ISP. Thus, from an ex ante perspective the expected profitability of innovation investments decreases for the smaller content provider in comparison to the case where both types of content providers win the exclusivity with 50% chance. This might lower the incentive for smaller content provider to innovate and might be inefficient. Note, however, that if lower administrative costs of the larger content provider are the reason for adoption of the innovation of the larger provider, than this appears to be more efficient than a simple 50% chance adoption rule.

To the first, the standard theory on exclusionary behavior applies. According to the Chicago Critique, the opposite side of the market, where the exclusivity leads to potential exclusion, has no incentive to sign a contract which excludes a more efficient business partner. Again, there are a number of prominent exceptions. For example, it might be of importance whether one market side is fragmented, that is, whether there are a number of small innovative service providers which cannot coordinate their behavior. In such a situation a dominant ISP may collect a selection of exclusive contracts with a critical mass of content providers, effectively foreclosing rivaling ISPs from this segment. This does not appear to be the right description of the situation, but in the future a larger number of providers of innovative services might appear. However, each type of innovative service might be considered as a separate market. Furthermore, the economics of the Internet (due to network effects) support larger, and not a fragmented mass of smaller, content provider. So in this sense the set up would not resemble the conditions needed in order to run a fragmentation argument in the exclusive dealing context. This example illustrates that it is important to not generally prohibit exclusive content on the basis of anti-competitive concerns. Instead a case-by-case approach might be more appropriate.

To the second sub-question, it appears generally difficult to deter entry into the provision of best effort network with the help of exclusive content on the innovative services. The level of competition between ISPs on the current best effort network appears to be high. In order to deter entry in the best effort, ISPs might use some bundling pricing strategies whereby they would try to leverage some power held in one market into another more competitive market.[150] Again, the Chicago Critique would argue that there is only one monopoly profit: while a monopolist could use monopoly profits in one market ("innovative services") in order to subsidize the product in another market ("best effort"), the monopolist would never get a higher effective price for the competitive product ("best effort") than its production costs in the competitive market.

There are a number of papers which challenge this notion, arguing in particular that economies of scale and non-perfect competition can imply that entry deterrence of more efficient competitors is profitable.[151] However, a number of

---

[150] Bundling occurs when one company offers two products A and B and product A is only available in a bundle with product B whereas product B is also available outside the bundle.

[151] For example Whinston (1990) finds that with fixed sunk costs of entry, bundling the product with a monopoly product can prevent entry. The idea is that the monopolist can credibly threaten to subsidize the product with revenues from the monopoly market in case of entry. So the monopolist uses the bundling as commitment to aggressive pricing in

those papers assume that bundling two products is credible in the sense that it is impossible to market the two products independently once the bundling decision has been undertaken. This is a feature which is present in some markets, for example when the bundled product is technologically incorporated in the other product. In the context of the best effort Internet and innovative services, it is conceivable that an ISP might try to bundle the two products together: while there is a natural connection between the two types of products due to the fact that they run on the same infrastructure, it is unclear how credible this bundling decision is and whether the IPS would not offer its monopoly innovative service on top of other operators' best effort Internet service. Again, the observation that we would generally not expect profitable exclusionary behavior when one market is competitive calls for a case-by-case inspection of potential exclusive content agreements.

Contrary to the above discussion, vertical agreements like exclusive dealing are often used in order to generate efficiencies. Typically, vertical agreements can address issues of double marginalization, moral hazard or hold-up problems. In the present context one might argue that exclusive content agreements might help ameliorate the negative effects of an inherent hold-up problem.[152] Also in highly dynamic markets exclusivity agreements might enable market entry and might be abandoned shortly after successful entry as in the case of the Apple's iPhone.

The "Best Effort Plus" business model disrupts the industry standard of **interoperability** and offers specialized services only on proprietary networks. This implies that network effects might be limited for those innovative services which rely on network effects (e.g., interactive video gaming applications or e-Health applications with the objective to combine specialist knowledge from a broad base of doctors).

The **risk of foreclosure** was discussed in detail under the question of the ability and incentive to exclude rivals through exclusivity and bundling strategies. It should be noted that these issues are unlikely to arise in a situation where even the innovative segment is regulated so that no discrimination is allowed.

---

the contested market. Another approach shows that bundling two markets where entry is possible in both markets makes entry in either one of the markets less likely, see Nalebuff (2004).

152 A hold-up situation occurs when it would be mutually beneficial for two players to cooperate in order to provide a certain good or service, but they refrain from doing so out of fear that this cooperative behavior could be exploited by the other party, ex post.

In contrast to foreclosure concerns discussed under the "Congestion-Based Model", this type of foreclosure does not assume that the ISPs are vertically integrated into content. Foreclosure concerns which require vertical integration as a prerequisite can, however, also apply within this setting. Apart from exclusivity and bundling techniques, in this model the ISP might also try to discriminate within the best effort network with the help of network management techniques. However, this does not constitute a change in comparison to the status quo. What does constitute a change, however, is that a vertically-integrated ISP might try to engage in discriminatory pricing of higher quality classes. Concerning foreclosure risks on the best effort network, the previous limiting comments apply: access regulation limits concentration in Europe, the level of vertical integration is moderate compared to the US and it is likely that non-discriminatory regulation is going to be effective for the best effort network. Furthermore, precedent on such behavior is thin so far.

As in the status quo, one major requirement for functioning competition in the provision of Internet services to be effective relates to the level of transparency of the use of network management tools for end users and content providers.

Table 12 summarizes the effects. The take-away points are: ISPs gain the option to offer premium services to these content providers who need their content delivered at a premium rate (value added service). Minimum guaranteed reserved bandwidth for priority novel services would ensure their quality or even viability. However, the risk of foreclosure due to exclusive agreements and bundling strategies might be increased, which might not be a major concern with access regulation in the European environment.

Table 11: Impact assessment of Best Effort Plus (BM 1)

		Possible impact assessment
Congestion		Persistent congestion on best effort Congestion levels compatible with agreed quality levels Limited spillover of congestion from high-quality network to best effort network
End users	Prices	Presumably unchanged for best effort Independently marketed innovative services
	Participation	No change for best effort Participation in innovative services
Content providers	Prices	Presumably unchanged for best effort Increased prices for premium class reflecting higher quality
	Participation/ Innovation	Limited effects for best effort class Premium class encourages creation of new premium services
ISPs' incentives to invest		Increased incentives to invest in capacity for paid services No (limited) incentive to under-invest in order to achieve higher prices for high quality
Regulatory costs	Ex ante intervention	Transparency for content providers and end users wrt network management tools (as in status quo). Regulation or industry commitment necessary to define boundaries of innovative services. Substantial ongoing regulatory oversight costs necessary to verify that the services fit within the defined class boundaries.
	Ex post intervention (competition concerns)	Increase in the risk of foreclosure by dominant vertically integrated ISPs due to potential discriminatory behavior with respect to access and conditions of high-quality class as well as exclusivity and bundling concerns between best effort services and innovative services might emerge.
Interoperability		No impact on best effort interoperability. Innovative services potentially fragmented.

Source: ESMT CA.

4.2.3. BM 2: Quality Classes: Content Pays

The "Quality Classes – Content pays" business model offers different quality classes open for every application. In turn, content providers pay a premium for higher quality transmission. However, end users are still on a uniform flat rate in this model and experience the quality as chosen by the content provider.[153]

This model preserves the interoperability between different interconnecting ISPs in the sense that every bit of traffic is transported on the best effort level. To the extent that interoperability on best effort preserves the non-excludability of the current Internet structure, we expect that a congestion problem persists due to the **common-pool** characteristics. In fact, congestion problems on the best effort level might be aggravated in the short term given that higher quality traffic gets priority on the same infrastructure. Even if this would be the case, for higher value traffic congestion problems are ameliorated and it is likely that overall welfare increases.

A number of economics papers focus on the effects of quality differentiation and its effects on congestion given the level of infrastructure as well as the incentives to invest in infrastructure.[154] In summary, those models find that offering priority access is likely to result in an increase of total welfare **in the short run** (on a given level of infrastructure). This can be traced back to different types of efficiencies: (1) congestion is more efficiently allocated according to the sensitivity to delay of different types of content, (2) more profitable content gets priority, which induces end users to switch to content provider or (3) participation is increased with differentiated access. However, it is true that the service level in the best effort class (lowest quality class) is likely to decrease once a priority access is introduced: whereas congestion in the priority class decreases, congestion is likely to increase in the best effort.

Under the “Quality Classes – Content Pays” business model, ISPs are able to offer **different quality classes** to all applications. According to the economic first principles, product and price differentiation is likely beneficial in terms of total welfare. In this situation, content providers are enabled to provide higher quality services to their end customers using higher quality transmission levels. In the counterfactual where congestion problems persist on the best effort infrastructure, it can be expected that smaller content providers offering delay sensitive content are crowded out from the infrastructure. Crowding out affects especially smaller content providers for two reasons: first, they might not have the financial means to

---

153 For simplicity, this business model assumes that there is only one access flat fee for an end user which is not differentiated according to bandwidth. Accordingly, we assume that each end user is able to receive high-quality traffic. The distinction of who pays for the higher quality classes is made in order to separate effects of the payment structure. It is, of course, conceivable that a mixture of the two business models BM 2 and BM 3 could arise in practice.

154 A detailed discussion of those papers can be found in the Appendix.

vertically integrate into the infrastructure like other larger content providers have (e.g., Google). Second, operators of content delivery networks like Akamai currently do not offer their services to smaller content providers. Thus, smaller content providers might have no access at all to higher quality transmission in contrast to large content providers in the current framework.

In the long run, an ISP might have an incentive to degrade quality in the low effort class in order to get an attractive price in the priority class in the sense that it under-invests into infrastructure expansion in comparison to the investment levels with a uniform quality level. In the previous business model, this was not a concern as content providers were not allowed to choose the quality classes for their services. Instead, the choice of quality regime was exogenously given depending on the type of content offered. To the contrary, in this business model content providers can choose the level of quality themselves. There are no robust theoretical results in relation to this: some papers show that incentives to invest increase, some that they decrease in the presence of quality differentiation. One major drawback in this literature is that most papers deal with monopolistic platforms and disregard the effects of platform competition entirely. The economic first principles show, however, that the effect of competition on investment incentives can be substantial.

As a remedy to potential problems of strategic degradation of quality the introduction of a minimum quality of service regulation has been discussed. Such a regulation might be helpful in certain circumstances, but does also bear its own risks. We are not aware of an economic paper which models the effects of minimum-quality regulation on the resulting offered quality levels. However, a few more general remarks can be made. First of all, socially optimal quality differentiation is likely to decrease quality in the low class in comparison to a uniform quality level. Thus, a regulation which applies the simple rule that the best effort quality level as it stands today should be guaranteed is likely to be overambitious. Consequently, setting the right minimum-quality standard might be a very difficult task for the regulator.

Second, a minimum-quality regulation would introduce a constrained optimization problem for the platform. For ease of illustration, consider a monopolistic platform and a situation in which the regulator sets the minimum quality to the socially optimal level for the low-quality class (i.e., below the uniform quality). The monopolistic ISP still has the same incentive to shape its price and quality structure such that it can extract the maximum rent from the market participants. In order to incentivize content providers with high willingness to pay, it could for example increase the price for the low-quality type: if the regulation ensures that the optimal level of quality is provided for content providers requiring low transmission quality and the ISP reacts by increasing the price for this quality level, then the social welfare implication of the minimum quality of standard regulation are unclear.[155] In this sense, it is not clear to what extent a minimum-quality regulation would increase social welfare given that the potential incentive to degrade quality persists.

In this business model, the content provider pays for higher qualities – on average this might imply that the price to content providers increases in comparison to the current situation. According to the **waterbed effect** in two-sided markets, this increase is likely to decrease the price paid by end users. This in turn might imply an increase in the participation of end users.

Finally, we discuss the issue of **interoperability** of quality classes. The welfare effects of various ISPs potentially choosing different quality levels have not been discussed in the economic literature. This seems natural since most papers consider monopolistic ISPs. The exception here is the model by Hermalin and Katz which also extends to duopoly platforms. The paper finds that the unique equilibrium with single-homing end users is a symmetric equilibrium where each ISP provides the same quality class. This implies that interoperability is not an issue. In contrast, the paper by Njoroge et al (2010) shows that if ISPs are allowed to choose a uniform quality level only, they are going to differentiate in equilibrium.[156] This paper makes the assumption that the transmission quality of an end-to-end connection is determined by the lowest of the quality choices of the involved platforms (i.e., the access platform to the end users and the one to the content provider). However, the authors do not consider whether it is socially desirable to have ISPs invest in different quality levels or whether one should encourage similar qualities across platforms.

These questions relate to the trade-off covered in the section on Economic first principles: on the one hand, consumers with heterogeneous preferences might be better off if they are able to choose from a wider range of options. On the other hand, individual utility is limited if the quality choices between networks are not

---

[155] Alternatively, the ISP might also decrease the price/increase the quality for the high quality. In the classical screening model, the high quality level would be set at the socially optimal level (see argumentation above). In the two-type model, this implies that there is no use in increasing the quality for the high type. However, with a continuum of agents, it might also be worthwhile to increase quality for the high types. The social welfare implications of the regulation in this setting are also ambiguous.

[156] The paper shows that platforms are going to differentiate maximally in terms of quality if no termination charges are allowed (e.g., specific access charges to content providers which are not directly connected to the platform). However, quality differentiation also persists in the case where termination is allowed.

compatible and no network effects (on quality) can be reaped. There might be a few specific arguments in favor of quality standardization on the Internet:

- **Separation of content and consumer networks**: The large degree of separation between content and consumer networks implies that a large share of the traffic will have to be delivered from one network into another. If the quality class for the end-to-end connection is determined by the lowest quality class, then this essentially means that qualities classes should converge between different networks.
- **Technical reasons**: To ensure quality of delivery, packets of a different quality need to be somehow identifiable (e.g., as a part of a packet header). Ensuring that this identification is understandable throughout the Internet requires some form of protocol standardization and thus the potential convergence of quality classes.
- **Transaction costs**: The desire to reduce transaction costs might lead to a reduction in the number of quality classes overall. If this happens, then lower bilateral bargaining costs among ISPs would be a likely consequence. However, it might also be associated with lower implementation costs.

There are a number of ways to reach interoperability of quality standards. They could emerge in equilibrium (as for example in the model by Hermalin and Katz), be agreed upon within international industry organizations or be set by some international regulatory body.

Within this business model the **risk of foreclosure** might be increased in comparison to the current status quo as the ability to charge for higher quality transmission opens a further source of uncertainty. This uncertainty could be used in order to charge discriminatory prices to rivals in order to try to squeeze them out of the market. The risk of foreclosure can be mitigated if the ISP can credibly commit to non-discriminatory access. Even though transparency might help create such commitment, it is uncertain if commitment is feasible. Furthermore, network management techniques might be used to also discriminate within the best effort network as in the current status quo.

Concerning foreclosure risks, the previous limiting comments apply: access regulation limits concentration in Europe, the level of vertical integration is moderate compared to the US and it is likely that non-discriminatory regulation is going to be effective for the best effort network. Furthermore, precedent on such behavior is thin so far.

As in the status quo, one major requirement for functioning competition in the provision of Internet services to be effective relates to the level of transparency over the use of the network management tool for end users and content providers.

Table 13 summarizes the effects. The take-away points are that higher qualities facilitate new content. Charging content providers rather than users for the higher quality levels is likely to maximize the value of the platform. However, the model introduces the risk of under-investment into the infrastructure due to a strategic incentive: degrading quality in best effort might hike up the price for higher quality levels. If this proves to be problematic, it might be addressed with a minimum quality of standard regulation which in turn bears its own risks.

Table 12: Impact assessment of Quality Classes – Content Pays (BM 2)

		Possible impact assessment
Congestion		Reduced congestion for high-quality premium paid services Likely increased congestion for the low-quality best effort class
End users	**Prices**	Little change relative to benchmark (status quo) Possibly somewhat lower prices due to the waterbed effect
	Participation	Increase due to availability of premium services and possibly lower prices
Content providers	**Prices**	Increased prices for premium class (reflecting higher quality)
	Participation /Innovation	Ambiguous: Premium class encourages creation of new premium services Potential decrease in innovative activity due to rent extraction (waterbed effect) in particular of lower quality services
ISPs' incentives to invest		Ambiguous: Increased incentives to invest in capacity enabling new premium content At the same time potential incentive to degrade best effort quality
Regulatory costs	**Ex ante intervention**	Transparency for content providers and end users wrt network management tools (as in status quo). Potential introduction of minimum quality of service to limit eventual negative feedback effect of higher qualities on best effort.
	Ex post intervention (competition concerns)	Increase in the risk of foreclosure by dominant vertically-integrated ISPs due to discriminatory behavior with respect to access and conditions of quality classes.
Interoperability		Fragmentation risk concerning higher quality classes if industry-wide standards over quality classes would not emerge.

Source: ESMT CA.

4.2.4. BM 3: Quality Classes: User Pays

The “Quality Classes – User pays” business model offers multiple quality classes for users that are designed to match the different usage patterns of different end users. For example, end users who frequently use interactive applications might choose the quality class which is more suitable for dealing with such applications, that is, that offers a low level of delay and jitter. Other users, who focus on multimedia applications, might choose another quality offering characterized by low packet loss and high bandwidth, and so on.[157]

The introduction of multiple flat fee packages could help in addressing some of the **common-pool** resource problems, but only to a very limited degree. Flat-fee pricing structures in general encourage excessive usage and thus contribute to congestion.[158] So the problem of overuse and congestion would remain largely unsolved with flat rates. However, introducing quality classes for end users also allows the prioritization of traffic and thus allows existing infrastructure to be used more efficiently (see discussion of BM 2 for more detail on this).[159]

Multiple quality classes with differentiated flat fees would also introduce some **product and price differentiation** on the end user side of the market. In general, such developments can be beneficial for (most of) the end users, content providers, and ISPs as it increases product variety. For end users, some welfare gains could be generated because users could choose options aligned more closely with their preferences. Additionally, the emergence of low-quality packets at low prices would also allow broadband access to subscribers who were priced out of the

---

[157] The distinction of who pays for the higher quality classes is made in order to separate the effects of the payment structure. It is of course conceivable that a mixture of the two business models BM 2 and BM 3 could arise in practice.

[158] By introducing a menu of different options, end users will self-select into different packages according to their preferences and expected usage patterns. To the extent that high-volume users self-select into high-quality classes, some form of volume-related pricing emerges. For example, a user who downloads a lot of content would likely choose a more expensive option than an occasional light user. In this way, the subsidy from the low-quantity users to high-volume users would be somewhat reduced but not completely eliminated as in volume-based pricing.

[159] Even though one could argue that end users are actually the ones initiating most of the traffic, in most of the cases they are unaware of the volume of traffic they are generating. This implies that they are also unaware of the quality of transmission necessary to ensure that the quality of experience at the end user side is sufficient. Although it appears reasonable to assume that consumers can be informed about the volume of traffic implications of different applications as the example of mobile Internet users shows, it is one complicated step further to educate consumers about the level of service needed to ensure a good quality of experience of different applications. To date it might be difficult for end users to choose technical quality classes on the basis of their user profiles (mix of consumed services). To the extent that end users find it difficult to choose the appropriate quality class, it might be difficult to increase the efficiency of capacity utilization by offering quality classes to the end user.

market at a uniform price. Therefore, this differentiated pricing strategy would also have a desirable effect of broadening the end users' base.

In terms of distribution of rents, product and price differentiation would reduce the amount of cross-subsidies between different user groups. So high volume users might be worse off relative to uniform prices because they would likely self-select into a more expensive premium package while on the other hand, prices for occasional low volume users would likely decrease, relative to a uniform price (see discussion of BM 0 for more detail on this).

Even though content providers could not choose a specific quality class for their services, it is likely that the quality of services they offer to the end users would increase, especially in the upper segment. This is because users who would be interested in their services would be expected to choose a quality class appropriate for the services (otherwise the service would not be of sufficient quality or would simply be unavailable).

In comparison to the previous business model, the end user pays for higher qualities. Thus, the **waterbed effect** is likely to favor content providers: Other things equal, price differentiation allows the platform owner to extract surplus from end users more easily. Because of this, the platform owner will often want to alter its pricing strategy for content providers in order to increase participation and thereby boost valuations for the end user. Thus, with differentiated prices on the end user side, the ISPs would likely have incentives to lower fees on the content side. As a consequence, end users would potentially receive a broader Internet offer for their higher prices. However, it is unclear if prices to the other side of the market can be fully adjusted. For example, there are no proper prices for content delivered in peering agreements between ISPs, so the prices cannot be lowered. Accordingly, the waterbed effect may not work fully in this business model.

Assuming that end users' demand is more price elastic than content providers', and considering the fact that end users become more and more valuable for the other market side, two-sided market theory predicts that content providers' prices should be subsidizing the price of end users in order to entice their participation in the social optimum. In this sense, pricing end users for higher quality and thus demanding higher average prices is likely to decrease social welfare in comparison to the case where content providers' are priced for the higher quality.[160]

---

[160] In relation to price discrimination the following should be noted: the greatest gains from price differentiation in two-sided markets arise if differentiated prices can be offered

Introduction of multiple quality classes would likely increase **incentives to invest and innovate** both for ISPs and content providers. It is unlikely that existing infrastructure would support multiple quality classes, so ISPs would likely need to undertake significant investments in additional capacity to enable the quality classes they designed. Just as in the previous business model, the question however emerges whether ISPs might have an incentive to invest less into the infrastructure in order to generate higher prices for higher quality than in a situation where no quality differentiation is allowed. However, as pointed out in the section on BM 2, there are no robust theoretical results on this. Furthermore, the incentives to degrade lower qualities when end users are paying is likely lower as the willingness to pay of (private) end users for higher quality is likely to be limited.

In terms of the investment incentives of content providers, it seems that the emergence of quality classes could encourage them to provide additional, innovative services unavailable with purely “best effort” internet. For example if sufficient bandwidth capacity were “set aside” or allocated for video download services for premium end users, more high quality downloads (e.g., HD movies) would be possible than in pure best effort regime. In general, the impact of additional quality classes on content providers would be positive.

Introduction of multiple quality classes for end users could be somewhat problematic from the **interoperability** point of view. Clearly, because of the technical requirements of different quality classes not all content would be available to all end users, which would reduce the positive impact of network effects. For example, end users who would subscribe to the “basic” quality class could not reasonably have access to HD video streaming capabilities. However, all users would still have access to the basic services, so the negative effect of the interoperability would be felt most severely in the highest quality classes. However, because such services are not available under the current basic structure, welfare losses due imperfect interoperability are only relative to the hypothetical first best scenario. In comparison to the status quo, interoperability problems of higher quality levels should be of minor concern as in the analyzed

(are allowed) on both sides of the market. If differentiated prices are allowed only on one side of the market, it is usually more beneficial to offer differentiated prices on the side of the market in which the participation of buyers with low willingness to pay provides more positive externalities. Keeping average prices constant, it might thus be more favorable in the future to price discriminate the end user side in order to stimulate participation.

business model higher qualities are available to at least a subset of end users as opposed to no consumers at all.[161]

When end users pay for higher quality classes, no additional **risk of foreclosure** is introduced into the pricing structure in comparison to the current status quo. However, reasonable network management practices could also be abused in this setting to try to exclude rival content. As pointed out previously, such strategic activities are less likely in Europe than in the US and even there only limited precedent exists.

As in the status quo, one major requirement for functioning competition in the provision of Internet services to be effective relates to the level of transparency over the use of the network management tool for end users and content providers.

Table 14 summarizes the effects. The take-away points are that also in this model higher qualities facilitate new content. However, charging users rather than content providers for the higher quality levels is likely to lead to lower value of the platform and lower incentives to invest in the platform than in the previous business model.

---

161 There are also other problems, related to the lack of standardization of quality classes. If there were no industry standards of ensuring quality levels commonly agreed upon, then from the end user's perspective, there would be no easy way to allocate the potential blame for an outage or reduced quality of service. Monitoring the quality of service is a difficult task, especially for end users. Without industry-wide standards, the ISPs could not actually guarantee a specific level of quality for a given quality class because they do not control the performance and data transfer outside of their own network. For example, if a low-latency end user tried to download some latency-sensitive content and experienced unacceptable delay, he would have no way to determine if this was caused by inadequate service at the content provider's ISP, interconnecting ISPs or his own eyeball ISP. Therefore, without Internet-wide interoperability, technological protocols and standardization of the quality classes, they would only have relative (as opposed to absolute) meaning, i.e., a subscriber to a premium "gold" package could expect a better performance of his Internet connection than a subscriber to an intermediate "silver" package, especially in periods of congestion, but the absolute level of the overall service could not be guaranteed. This effect would likely reduce the desirability of the premium packages, as they all (without standardization) would in some sense be "best effort" packages.

Table 13: Impact assessment of Quality Classes – User Pays (BM 3)

		Possible impact assessment
Congestion		Reduced congestion for high-quality premium paid services Likely increased congestion for the low-quality best effort class
End users	**Prices**	Ambiguous: Availability of inexpensive low-quality flat rates Increased flat rate for premium class (reflecting higher quality)
	Participation	Ambiguous: Increased participation due to inexpensive low-quality flat rates and availability of premium services Decreased participation due to potential price increases (waterbed effect)
Content providers	**Prices**	Little change relative to benchmark (status quo). Possibly somewhat lower prices due to the waterbed effect.
	Participation /Innovation	Premium class encourages creation of new premium services
ISPs' incentives to invest		Ambiguous: Increased incentives to invest in capacity enabling new premium content At the same time potential incentive to degrade best effort quality. Likely significantly lower incentives than under content pays business model due to lower profitability of platform
Regulator y costs	**Ex ante intervention**	Transparency for content providers and end users wrt network management tools (as in status quo)
	Ex post intervention (competition concerns)	No increase in the risk of foreclosure by dominant vertically-integrated ISPs
Interoperability		Fragmentation risk concerning higher quality classes if industry-wide standards over quality classes would not emerge

Source: ESMT CA.

4.3. Discussion of net neutrality definitions and their impact on possible market outcomes

The previous sections described important economic first principles relevant in the context of Internet business models and potential regulation. Those principles guided through the detailed analysis of economic effects of the various proposed business models. This section highlights the interaction between different business models and concepts of net neutrality discussed in the academic literature. There are two fundamental points of view with which one can analyze the business models on this background. First, some of the proposed business models incorporate some form of "commitment" in the sense that they restrict the ISPs behavior. In order to ensure a welfare-enhancing commitment rule, monitoring might be necessary to evaluate the need for further regulation. This angle relates essentially to the "Best Effort Plus" model where either an industry commitment or regulation is needed in order to delineate innovative services which can negotiate higher qualities from traditional services which are all delivered on a best effort level. Second, regulatory intervention in the form of the implementation of net neutrality rules might restrict or render some business models impossible. In the following, we focus on this second approach and we analyze the relationship between different forms of net neutrality and the various business models.

4.3.1. Connection of business models with definitions of net neutrality

Generally, the regulation of the Internet has been discussed under the term of net neutrality. No single and commonly agreed upon definition of the concept of "net neutrality" exists. In the academic discussion, net neutrality encompasses at least four different regulatory regimes, which differ with respect to allowed technological differentiation (prioritization) of some types of data as well as with respect to allowed pricing for prioritized data transmission.

- Definition 1: In the most extreme and ideological notion, net neutrality requires absolutely identical treatment of all internet data transmissions, both with respect to quality as well as price.[162] Since most ISPs engage in

[162] It is often argued that equal treatment of traffic is the only fair mechanism of traffic transmission. However, this depends on the notion of fairness. From a more abstract perspective a major distinction is whether one defines "equal treatment" in comparison to the status quo or relative to well-functioning competition. The latter definition allows discrimination based on application or technology-specific market failures. It could for example be argued that each application should have the same "quality of experience" level for the end user. This would imply that that very time-sensitive applications would

reasonable network management practices, resulting in partial prioritization of certain data types, this proposal restricts the industry beyond its current *status quo*.

- Definition 2: In the second possible meaning net neutrality would legally codify the current *status quo* as a binding regulation. This definition does not allow any price differentiation across different data packets, but would allow some differentiation of treatment of individual packets based on their type and required quality characteristics, which are currently often described as “reasonable network management” practices aimed at increasing the efficiency of the existing infrastructure.
- Definition 3: A third definition of network neutrality would allow differentiation of data streams both in terms of their quality characteristics as well as prices charged for transmission of data packets of different qualities, but which at the same time would require an open and non-discriminatory access to different price-quality tiers to all Internet participants.
- Definition 4: Finally, net neutrality definition may be expressed in terms of an absolute ban on any termination fees that eyeball ISPs could charge content providers. The concept of “termination fees” originates from telecommunications networks, where they describe payments the operator of the party originating the call must pay to the other telephone network operator to “terminate” the call, that is, to reach the intended recipient. Similarly, in the context of net neutrality debate “termination fees” denote payments that ISPs could charge certain content providers to reach the ISP’s subscribers. Failure to pay the fee would result in the ISP blocking the traffic from the non-paying site to its customers, and thus would prevent its own customers from reaching the non-paying sites. In the less extreme version, failure to pay the termination fee would not result in the complete block of traffic of the non-paying site but would instead result in the provision of a lower quality (priority of data transmission) relative to sites that would pay the termination fees.

Since different applications have different requirements regarding quality of data transmission (bandwidth, delay, jitter, etc), one can argue that ISPs have

---

be fairly prioritized in order to guarantee the same level of quality of experience as less time-sensitive applications. Briscoe (2007) notes that: “Fair allocation of rates between flows isn't based on any respected definition of fairness from philosophy or the social sciences. It has just gradually become the way things are done in networking. But it's actually self-referential dogma.”

legitimate reasons to treat different types of data differently in order to optimally utilize their available bandwidth capacity. These so-called "reasonable network management" practices are a feature of the Internet as it operates right now. For that reason, even the proponents of network neutrality regulation tend to agree that ISPs should have some flexibility to efficiently organize data traffic on their networks and ensure high-quality experience for their end users. This essentially rules out the first definition as a relevant option for regulation.

Table 15 illustrates the relationship between the different business models and definitions of net neutrality, which have been put forward by proponents of net neutrality in the public discussion, in particular in the US. They serve as a reference for the academic evaluation of the competitive effects for the stylized business models.[163] Table 15 shows that a number of net neutrality definitions would undermine different business models.

Table 14: Connection of BM with definitions of net neutralityutrality

	Def 1: Equal treatment	Def 2: Reasonable network management, but no payment for quality by CP (status quo)	Def 3: Non-discriminatory quality classes	Def 4: No termination fees
BM 0: Congestion-Based Model	Violated*	Not violated	Not violated	Not violated
BM 1: Best Effort Plus	Violated	Violated	Violated	Violated
BM 2: Quality Classes – Content Pays	Violated	Violated	Not violated	Not violated
BM 3: Quality Classes –User Pays	Violated	Not violated	Not violated	Not violated

Note: *Also the "Congestion-Based Model" assumes that reasonable network management is undertaken.

Source: ESMT CA.

[163] Another issue is whether the European regulatory framework would allow the implementation and enforcement of the definitions discussed in the US context. This is a legal question and is not assessed within this report.

As illustrated in table 15, the hypothetical implementation of different types of network neutrality definitions would render some of the business models infeasible:

- The “Congestion-Based Model” as well as the “Quality Classes – User Pays” business model are feasible under all net neutrality definitions with the exception of the strong interpretation embodied in definition 1 which would also not be compatible with the business practices in the current status quo.
- The “Quality Classes – Content Pays” requires a more unregulated environment, although it is consistent with some weaker definitions of net neutrality allowed in definitions 3 and 4.
- In contrast, “Best Effort Plus” implies some form of permanent regulatory exemption (holidays) for novel innovative services, and hence it is in principle not compatible with any form of net neutrality regulation in the broad sense, although the best effort component in BM1 is supposed to be regulated.

The above thus illustrates that the implementation of different forms of net neutrality impacts the above business models to a different extent. This demonstrates that any future regulation should carefully consider the impact on business models and the benefits which are foregone once certain models are restricted.

4.3.2. Implications of business model analysis on regulation

In order to draw some strong policy implications it would be necessary to rank different business models according to a set of criteria, such as total or consumer welfare. This represents a highly complex task as there is no single economic model underlying the different business models considered. Therefore the report does not attempt to rank the business models but instead discusses a number of trade-offs which the decision to regulate the Internet would likely involve.

The implementation of a **strong form of net neutrality** (definition 2) prevents “Best Effort Plus” and “Quality Classes – Content Pays”, but still allows the other two business models. This implies that some benefits of new business models can be reaped with net neutrality regulation whereas other efficiencies cannot materialize:

- **“Congestion-Based Model”**: Congestion-based pricing can be thought of as an extension of network management which is currently only achieved on a technological level. Therefore, the implementation of this model would bring some modest gains in the overall efficiency of the network. Moreover, the model would require little regulatory oversight and exhibit low risk with regards to anti-competitive behavior as compared to other business models considered. So the efficiency gains could potentially be realized at a low regulatory cost. However, this model does not allow for product differentiation and thus deprives end users and content providers of the benefits of this type of differentiation. In particular, it does not prevent that certain highly delay-sensitive content is crowded-out of the market at high congestion times. Therefore, the implementation of this model would bring at best only modest gains in overall efficiency of the network.

- **"Quality Classes – User Pays":** In comparison, the "Quality Classes – User Pays" model allows for different quality classes which opens possibilities for new content. However, charging users rather than the content provider for the higher quality levels is likely to lead to lower value and lower incentives to invest for the platform than in the business models "Best Effort Plus" and "Quality Classes – Content Pays". The regulatory risk related to foreclosure strategies seems smaller, though: the ability of a dominant ISP to favor a vertically-integrated content provider is lower.

It should be noted that a combination of the two business models would also be feasible under this type of net neutrality regulation. In this sense, the gains due to congestion pricing can be combined with efficiency gains achievable in other business models.

In contrast, the implementation of definition 3 or 4 would also enable the adoption of a business model which prices the content provider for higher qualities. Under the **"Quality Classes – Content Pays"** model new content is also facilitated via higher qualities – just as in "Quality Classes – User Pays". As regards comparison between the two models involving quality classes, the following trade-off applies: the increased risk of foreclosure where content pays for higher quality must be weighed against inefficiency related to pricing the consumer side. It is difficult to argue in general which of the two effects outweighs the other. One must take into account, however, that in the European environment foreclosure risks are mitigated by a number of factors. First, the ability to foreclose is limited by access pricing regulation prevalent in Europe. Second, the ISP's incentives to foreclose in Europe are also limited relative to, for example, the US because European ISPs seem in general less vertically-integrated into content than their American counterparts. Third, if consumer welfare rather than total welfare were to be used as the criterion for determining the desirability of the business models, then it is likely that the "Quality Classes – Content Pays" model would be preferable over "Quality Classes – User Pays" model. All these factors seem to favor to some extent the model where content pays (BM2) relative to where the user pays (BM3), although it is uncertain, if they are sufficient to mitigate the risk of foreclosure to a level where the first could be considered unambiguously preferred over the second.

Finally, under the "Best Effort Plus" model any net neutrality regulation could only apply to traditional services **not to novel innovative services**. This model would thus not prevail under any form of all-encompassing net neutrality regulation. Also under this approach to industry regulation, benefits from quality differentiation can be reaped for society. With content providers sharing parts of the additional infrastructure costs for new services with the ISPs, the end user may experience lower charges compared to the same offering under the status quo. However, it should be noted that this outcome might also be replicated with a combination of the two models where quality classes are allowed (BM2 and BM3).

Ultimately, the crucial comparison involves "Best Effort Plus" versus a combination of the quality classes models BM2 and BM3. The models are very different in their assumptions and consequently a clear delineation of the different risks and benefits is highly complex. Both models tend to increase the participation of end users and both open the way for content demanding higher quality of service. But it is beyond the scope of this study to assess which model leads to more end user participation, to more content innovation and/or to more infrastructure investment. In the longer run, the gap between the two models might partially decrease as the significance of the traditional services in the "Best Effort Plus" model might gradually diminish and unregulated novel services determined by commercial contracts and market forces will dominate the future Internet. However, to the extent that the "Best Effort Plus" model is characterized by exclusive agreements, the difference persists.

Appendix

Economics literature on quality classes

A number of papers focus on the effects of quality differentiation and its effects on congestion given the level of infrastructure. For example, congestion effects are modeled in the paper by Krämer and Wiewiorra (2010) using a monopoly platform as a comparison benchmark. The authors model net neutrality as zero additional charges (apart from general internet connection) to content providers. In contrast, the reference scenario allows ISP to offer priority access for an extra charge. In the short run – taking infrastructure as given – they find that allowing offering priority access decreases congestion for priority access clients in comparison to the net neutrality regime, but increases congestion for basic service (best-effort) clients.[164] Total welfare increases in the short run as congestion is more efficiently managed.[165] End users' surplus does not change as the average waiting time is equal in the two regimes. However, the increased profits of content providers are appropriated by the extra charge of ISPs. Thus, ISPs have an incentive to engage in quality differentiation in the short run since it increases their profits.

A similar effect is found in the paper by Cheng et al (2010). In comparison to Krämer and Wiewiorra who analyze a continuum of content providers that are differentiated according to the delay sensitivity of the content, Cheng et al. consider the case of two content providers with different profit margins which are situated at the ends of a Hotelling location model. They find that either only the high-margin content provider chooses priority access or both content providers choose it. The authors point out that no content provider is effectively prioritized in the latter case.[166] Moreover, the ability to offer priority access merely implies that surplus is shifted from content providers to the ISP in this case. End users' surplus is unaffected. However, in the case where only the high-margin content provider buys the priority access, total welfare increases as this equilibrium occurs precisely when it is socially beneficial, that is, when the difference in profit

164 In their model no content provider is expelled from the platform in the discrimination scenario as congestion-insensitive applications chose the non-priority access and do not pay an extra charge: Profits of those applications decrease as congestion increases. However, since there are no fixed costs of operating the service, this has no implications on the participation decision of the content provider.

165 Congestion requires some form of rationing and economic theory predicts that introduction of pricing is often an efficient form of rationing.

166 This situation represents a type of prisoner's dilemma.

margins is sufficiently large. However, in this situation some end users are worse off as the faster connection induces them to choose the higher margin content provider – who is further away from their preferences.

Choi and Kim (2010) present another modeling approach which is very similar to Cheng et al. The main difference lies in the way both papers treat the sale of priority access. Whereas Cheng et al. assume that content providers pay a fixed price per data packet, Choi and Kim assume that the exclusive right to priority access is sold in an auction to the highest bidder. This set up does not seem to depict the proposed business model. However, the two papers illustrate that the results are rather sensitive to the modeling assumptions as Choi and Kim derive strikingly different outcomes in the short run. In their model only one content provider – the high margin content provider – is going to acquire priority access. This implies some consumers will switch to the less-preferred content providers. This loss in welfare is only compensated for if the difference in profit margins of content providers is sufficiently large. Otherwise total welfare decreases in the short run.[167]

However, quality differentiation might also give rise to a strategic incentive of the ISP to degrade the best effort quality in order to urge content providers to upgrade to higher quality classes. In the previous business model, this was not a concern as content providers were not allowed to choose the quality classes for their services. Instead, the choice of quality regime was exogenously given depending on the type of content offered. To the contrary, in this business model content providers can choose the level of quality themselves.

The effects of second-degree price discrimination based on quality differentiation are less clear cut than in the case of second-degree price discrimination based on volume discounts (see principle 2 und 3 in the previous section). Effects depend among other things on which uniform quality is chosen when second degree price discrimination is not feasible. Hermalin and Katz (2007) provide a model which

---

[167] Another paper by Jamison and Hauge (2008) points out that allowing for priority access might in fact encourage the participation of smaller content providers who would access the network on premium services. Accessing the network on premium service implies that the lower type content providers can compensate with speed for the fact that they only provide a low content value for consumers. Abstracting from the fact that the monopoly platform might have an incentive to degrade quality in the low-quality class, the authors find that introducing quality differentiation increases participation as well as investment into capacity.

analyzes the strategic effects of quality degradation.[168] In this model, quality is a direct choice variable of the ISP, so the operator offers a menu of prices for different quality classes. The authors show that the general result of screening models applies whereby the ISP has an incentive to lower the quality of the low-quality class below the uniform quality level in order to induce the high type to choose the high-quality level.[169] In terms of welfare implications, the authors show that content providers with low valuation who would not participate with a uniform quality standard are induced to participate when a lower quality level is offered. This effect increases welfare in comparison to the uniform quality case. Moreover, content providers with higher valuation are better off with differentiated quality levels as they are able to purchase a more efficient quality level. However, there exist some intermediary types who purchase a lower quality under differentiated quality levels than under the uniform standard and for whom this represents a stronger deviation from the efficient quality choice. This effect reduces social welfare when differentiated quality levels are allowed. Overall, the welfare effects are thus ambiguous. However, Hermalin and Katz argue that in most of the cases allowing differentiated quality levels is beneficial to society.

The paper by Hermalin and Katz models **investment decisions** of the ISP by assuming a general cost function for different quality levels. By linking the quality level to different waiting times and thereby modeling congestion explicitly, a couple of papers shed some more light on the investment decisions of the ISP. These papers investigate in particular the question whether the ISP has an incentive to under-invest in the provision of (best effort) infrastructure in order to keep the priority access attractive for content providers. Krämer and Wiewiorra (2010) show that allowing for quality differentiation increases the incentive for ISPs to invest in infrastructure. This is due to the fact that ISPs earn higher profits as they additionally earn profits from priority access. Overall congestion levels are thus decreased and quality differentiation has a positive effect on welfare in the long run.

However, this model of a monopoly ISP neglects the effect of available capacities on the price for preferential access. The conclusions drawn on investment incentives for ISPs thus have to be interpreted with care. In contrast, Cheng et al. (2010) and Choi and Kim (2010) take this countervailing effect into account. In the model set up of Choi and Kim, the results on the incentive to invest are ambiguous

---

168 The papers by Krämer and Wiewiorra and by Cheng et al. model quality as the average waiting time using queuing theory. They do not explicitly comment on the strategic incentive to degrade quality in the short run.

169 This is shown for a monopoly as well as a duopoly platform.

and depend on the magnitude of the following opposing effects: on the one hand, an increase in the quality of experience for the end user due to capacity expansion enables the ISP to charge higher prices from end users; on the other hand, a capacity expansion decreases the value of the priority access and thus decreases the ISP's profit. Cheng et al. arrive at less ambiguous results as they find that a monopoly ISP invests in most cases less when it is able to differentiate quality. Furthermore, they show that a monopolist without the ability to differentiate quality invests to achieve a socially optimal level of capacity.[170]

All of these papers analyze a monopolistic ISP. The exception is the Hermalin and Katz paper which does not model platform quality in terms of congestion. This paper shows that the duopolist platforms choose the same quality scheme as the monopolist platform in the unique symmetric equilibrium. This is due to the fact that end users are single homers as this implies that both platforms constitute a monopoly in terms of access to their end users. In this sense, platform competition manifests itself mostly in lower prices for end users implying that in a duopoly situation platform profits are competed away to the extent that the platforms are not differentiated.[171]

[170] In the presence of quality differentiation the ISP either over- or under-invests into capacity, but mostly under-invests.

[171] Another paper which analyzes a duopoly situation is Njoroge et al (2010). However, they model a different notion of net neutrality than the one discussed in the previous papers which relate to the ability to price differentiate quality levels. Instead, Njoroge et al address the question whether it is socially beneficial that ISPs are allowed to charge a termination fee to content providers which are not directly connected to their platform when they can choose a quality level for their network independently from each other. They find that end users as well as content providers are better off under a business model with positive interconnection charges as this increases quality of transmission in comparison to a business model where no such charge is allowed. In particular, content providers are also better off as they can generate higher profits under this scenario due of the fact that they are able to generate higher advertisement revenues with higher Internet qualities.

Bibliography

Aghion, P., and P. Bolton (1987). Contracts as a barrier to entry. American Economic Review 77(3): 388–400.

Alexandrov, A., and J. Deb (2010). Price discrimination and investment incentives, March 12. https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=IIOC2010&paper_id=340 (accessed January 21, 2011).

Armstrong, M. (2006). Competition in two-sided markets. RAND Journal of Economics 37: 668–691.

AT Kearney (2010). Internet value chain economics. http://www.atkearney.com/index.php/Publications/internet-value-chain-economics.html (accessed January 21, 2011)

Bagwell, K. (2007). The economic analysis of advertising. In *Handbook of industrial organization*, ed. M. Armstrong and R. Porter, vol 3, 1701–1844. Amsterdam: Elsevier.

Bauer, S., D. Clark, and W. Lehr (2009). The evolution of internet congestion. http://mitas.csail.mit.edu/papers/Bauer_Clark_Lehr_2009.pdf (accessed January 21, 2011).

Baumol, W.J., and D.G. Swanson (2003). The new economy and ubiquitous competitive price discrimination: Identifying defensible criteria of market power. http://www.abanet.org/antitrust/at-committees/at-pdiscr/special-resources/pdf/05_AB4089_UBAU_661–685.pdf (accessed January 21, 2011).

Bedre-Defolie, O., and E. Calvano (2010). *Pricing payment cards*. ECB Working Paper No. 1139.

Bernbom, G. (2000). Analyzing the Internet as a common pool resource: The problem of network congestion, International Association for the Study of Common Property. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.119.9942 (accessed January 21, 2011).

BMWI (2010). Netzneutralität – 11 Thesen für eine gesellschaftspolitische Diskussion. December 7. www.bmwi.de/BMWi/.../it-gipfel-2010-netzneutralitaet,property=pdf,bereich=bmwi,sprache=de,rwb=true.pdf (accessed January 21, 2011).

Briscoe, B. (2007). Flow rate fairness: Dismantling a religion. *ACM SIGCOMM Computer Communication Review* 37(2): 63–74.

Buyya R., M. Pathan and A. Vakali (2008). *Content delivery networks*. Berlin: Springer.

Cabral, L. (2008). Umbrella branding with imperfect observability and moral hazard. I*nternational Journal of Industrial Organization* 27: 206–213.

Cheng, H.K., S. Bandyopadhyay, and H. Guo (Forthcoming). The debate on net neutrality: A policy perspective. Information Systems Research.

Choi, J.P., and B.C. Kim (Forthcoming). Net neutrality and investment incentives. *RAND Journal of Economics*.

Crew, M., C. Fernando and P. Kleindorfer (1995). The theory of peak-load pricing: A survey. *Journal of Regulatory Economics* 8(3): 215–248.

Economides, N., and J. Tag (2007). *Net neutrality on the Internet: A two-sided market analysis*. Working Paper No. 07-27, Stern School of Business, New York University.

FCC (2009). Notice of proposed rulemaking. October 22.

FCC (2010). Connecting America: The National Broadband Plan. March 16.

Fudenberg, D., and J. Tirole (1984). The fat-cat effect, the puppy-dog ploy, and the lean and hungry. American Economic Review 74(2): 361–366.

Gill, P., M. Arlitt, Z. Li, and A. Mahanti (2008). The flattening Internet topology: Natural evolution, unsightly barnacles or contrived collapse? http://www.hpl.hp.com/techreports/2008/HPL-2008-47.pdf. (accessed January 21, 2011).

Hermalin, B.E. and M.L. Katz (2007). The economics of product-line restrictions with an application to the network neutrality debate. Information Economics and Policy 19: 215–248.

Hess, C., and E. Ostrom (2003). Ideas, artifacts, and facilities: Information as a common-pool resource. *Law and Contemporary Problems* 66: 111–145.

Houle, J.D., K.K. Ramakrishnan, R. Sadhvani, M. Yuksel, S. Kalyanaraman (2007). The evolving Internet: Traffic, engineering, and roles. http://www.cse.unr.edu/~yuksem/my-papers/2007-tprc.pdf (accessed January 21, 2011).

Jamison, M.A., and J.A. Hauge (2008). Getting what you pay for: Analyzing the net neutrality debate. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1081690 (accessed January 21, 2011).

Klemperer, J., and P. Farrell (2008). Coordination and lock-in: Competition with switching costs and network effects. In *Handbook of industrial organization*, ed. M. Armstrong and R. Porter, vol 3, 1970–2056. Amsterdam: Elsevier.

Krämer, J., and L. Wiewiorra (2010). Network neutrality and congestion-sensitive content providers: Implications for service innovation, broadband investment and regulation. Working Paper, Karlsruhe Institute of Technology.

Kruse, J. (2009) *Priority and Internet quality*. Discussion Paper No. 96, Helmut Schmidt Universität Hamburg.

Lee, R.S., and T. Wu (2009). Subsidizing creativity through network design: Zero pricing and net neutrality. *Journal of Economic Perspectives* 23: 61–76.

Levine, M.E. (2002). Price discrimination without market power. *Yale Journal on Regulation* 19: 25–26.

Lowry, T. (2009). Time Warner cable expands Internet usage pricing. *BusinessWeek*, March 31. http://www.businessweek.com/technology/content/mar2009/tc20090331_726397 (accessed January 21, 2011).

Mankiw, N., and M. Taylor (2010). *Economics*. Florence, KY: Cengage Learning.

MRG, Inc. (2010). IPTV global forecast, 2010 to 2014. http://www.mrgco.com/iptv/gf0610.html (accessed January 21, 2011).

Musacchio, J., G. Schwartz, and J. Walrand (2009). A two-sided market analysis of provider investment incentives with an application to the net-neutrality issue. *Review of Network Economics* 8: 22–39.

Nalebuff, B. (2004). Bundling as an entry barrier. *Quarterly Journal of Economics* 119(1): 159–187.

Njoroge, P., A. Ozdaglar, N.E. Stier-Moses, and G. Weintraub (2010). *Investment in two sided markets and the net neutrality debate*. Decision, Risk and Operations Working Papers Series, Columbia Business School.

OECD (2006). Internet traffic exchange: Market development and measurement of growth. April 5, 2006. http://www.ictregulationtoolkit.org/en/Publication.3081.html (accessed January 21, 2011).

OECD (2009). OECD communications outlook 2009. August 10. http://www.oecd.org/document/44/0,3343,en_2649_34225_43435308_1_1_1_1,00.html (accessed January 21, 2011).

Ovum (2010). Fall 2010 global Internet phenomena report. October 20. http://www.ovumkc.com (accessed January 21, 2011).

Pew Internet (2009). Home broadband adoption. June 17. http://www.pewinternet.org/Reports/2009/10-Home-Broadband-Adoption-2009.aspx (accessed January 21, 2011).

Phlips, L. (1983). *The economics of price discrimination*. Cambridge: Cambridge University Press.

Rasmusen, E., M. Ramseyer, and J. Wiley (1991). Naked exclusion. *American Economic Review* 81: 1137–1145.

Rochet, J.C., and J. Tirole (2006): Two-sided markets: A progress report. *RAND Journal of Economics* 37: 645–667.

Samanta, S.K., and H. Pan (2009). Pricing strategy of platform: An investigation to the Internet service provider (ISP) industry. Economics, Finance and Accounting Applied Research Working Paper Theory, Coventry University.

Sandvine (2010). Fall 2010 global Internet phenomena report. October 10. http://www.sandvine.com (accessed January 21, 2011).

Schuett, F. (2010). Network neutrality: A survey of the economic literature. *Review of Network Economics* 9(2): Article 1.

Segal, I.R., and M.D. Whinston (2000). Naked exclusion: Comment. *American Economic Review* 77: 296–309.

Spence, M. (1976). Product differentiation and welfare. *The American Economic Review*. 66(2): 407–414.

Tirole, J. (1988). *The theory of industrial organization*. Boston: MIT Press.

Weyl, E.G. (2010). A price theory of multi-sided platforms. American Economic Review. 100(4): 1642–1672.

Whinston, M.D. (1990). Tying foreclosure, and exclusion. American Economic Review. 80: 837–859.

Yoo, C.S. (2008). Network neutrality, consumers, and innovation. Scholarship at Penn Law Publication. http://lsr.nellco.org/upenn_wps/243/ (accessed January 21, 2011).

Yoo, C.S. (2010). Innovations in the Internet's architecture that challenge the status quo. Journal of Telecommunications and High Technology Law. 8: 79–100.

The authors

Hans Friederiszick is a managing director of ESMT Competition Analysis. He is an expert in applied microeconomics and industrial organization, specializing in competition economics. Dr. Friederiszick previously served on the Chief Economist Team at DG Competition for the EC where he focused on competition policy, antitrust investigations, merger control, and state aid.

Jakub Kałużny joined ESMT Competition Analysis in February 2009. He has extensive experience applying both empirical and theoretical economic analysis to antitrust litigation and regulatory matters. His areas of expertise include evaluation of the competitive effects of mergers, monopolization and vertical restraints, both in the United States and Europe. His analyses cover a wide range of industries including consumer goods, healthcare, agriculture, chemicals and information technology.

Simone Kohnz is a manager at ESMT Competition Analysis and has provided assistance and economic advice concerning mergers, cases of abuse of dominance, vertical and horizontal agreements and state aid questions. In particular, Simone Kohnz has worked on buyer power issues in merger analysis as well as abuse of dominance cases or sector inquiries. She has extensive experience with the analysis of rebate schemes and worked on a number of complex vertical issues.

Michał Grajek is associate professor at ESMT. Michał's research focuses on the frontier of industrial organization, where he utilizes advanced econometric methods to address issues related to modern economy. His area of research includes topics such as network effects, compatibility, and standardization with particular emphasis on Information and Communication Technologies. He has also worked on investment and regulation issues in network industries and standardization in international trade.

Lars-Hendrik Röller is President of ESMT European School of Management and Technology, Senior Advisor to ESMT Competition Analysis and a Senior Fellow at Bruegel. He served as the first Chief Competition Economist of the European Commission between 2003 and 2006.

About ESMT

ESMT European School of Management and Technology was founded in October 2002 by 25 leading global companies and institutions. The international business school offers Full-time MBA and Executive MBA programs, as well as executive education in the form of open enrollment and customized programs. The School also features in-house research-oriented consulting services. ESMT is a private university based in Berlin, Germany, with an additional location in Schloss Gracht near Cologne.

About ESMT Competition Analysis

ESMT Competition Analysis works on central topics in the field of competition policy and regulation. These include case-related work on European competition matters, for example, merger, antitrust, or state aid cases, economic analysis within regulatory procedures and studies for international organizations on competition policy issues. ESMT Competition Analysis applies rigorous economic thinking with a unique combination of creativity and robustness in order to meet the highest quality standards of international clients. As partner of the international business school ESMT European School of Management and Technology, Competition Analysis works closely together with ESMT professors and professionals on leading-edge research in industrial organization and quantitative methods.

More information:

ESMT Competition Analysis GmbH
Schlossplatz 1, 10178 Berlin
Phone: +49 (0) 30 212 31-7000
Fax: +49 (0) 30 212 31-7099
www.esmt.org/competition_analysis

ESMT
European School of Management and Technology
Faculty Publications
Schlossplatz 1
10178 Berlin

Phone: +49 (0) 30 21231-1279
publications@esmt.org
www.esmt.org