Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies

BACHELOR THESIS

Willingness to pay for online content in the Czech Republic

Author: **Michal Vala**

Supervisor: **Petr Polák, MSc.**

Academic Year: **2014/2015**

Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, May 15, 2015

_____

Signature

Acknowledgments

The author is grateful especially to Petr Polák his help and comments, and to his parents for the financial and other support, provided during his studies.

He would also like to thank PhDr. Tomáš Havránek Ph.D. for his LaTeX template that has helped substantially in the process of creating this thesis. The usual caveat applies.

Abstract

This thesis studies consumers' willingness to pay for online content in the Czech Republic. The study focuses on the newspaper industry in the online environment. Data analysed in the study come from a self-conducted online survey. Using the logit estimation method, following variables were identified to significantly affect the probability of having paying intentions—Gender, Occupation, Internet usage, News usage, Past online payment, and Past newspaper payment. Moreover, under no free newspapers scenario, a monthly subscription of printed newspapers was assigned with an average value of CZK 96.7, which is almost double when compared to the average of CZK 50.7 that would consumers pay for a monthly access to the digital edition of newspapers.

JEL Classification	D10, D12, L86
Keywords	WTP, online content, newspapers, Czech Republic
Author's e-mail	mic.vala@gmail.com
Supervisor's e-mail	polakpet@gmail.com

Abstrakt

Tato práce studuje ochotu spotřebitelů platit za online obsah v České republice. Studie se zaměřuje na novinový průmysl v online prostředí. Analyzovaná data byla získána s pomocí vlastního dotazníkového šetření. S použitím modelu logit byly identifikovány proměnné Pohlaví, Povolání, Používání internetu, Čtenost novinového obsahu, Předchozí platby na Internetu, a Předchozí nákupy novinových periodik, jakožto signifikantní ukazatele, které přímo ovlivňují pravděpodobnost, že čtenáři zaplatí za internetový obsah. Dále se ukázalo, že i v nepřítomnosti novinového obsahu, jež by byl zadarmo, měsíční předplatné tisku bylo ceněno v průměru na 96.7 Kč, což je téměř dvojnásobek průměrné vnímané ceny online předplatného, která činila 50.7 Kč.

Klasifikace JEL	D10, D12, L86
Klíčová slova	Ochota platit, online obsah, noviny, Česká republika
E-mail autora	mic.vala@gmail.com
E-mail vedoucího práce	polakpet@gmail.com

Contents

List of Tables

List of Figures

Acronyms

BDM Becker–DeGroot–Marschak

CDF Cumulative Distribution Function

LDV Limited Dependent Variable

LPM Linear Probability Model

MLE Maximum Likelihood Estimation

OLS Ordinary Least Squares

PBS Public Broadcasting Stations

WTA Willingness to accept

WTP Willingness to pay, willing to pay

Bachelor Thesis Proposal

Author	Michal Vala
Supervisor	Petr Polák, MSc.
Proposed topic	Willingness to pay for online content in the Czech Republic

Topic characteristics With the emergence of the Internet, the downfall of the traditional industries was often forecast. In contrast to these expectations, recent literature suggests that the Internet offers potential advantages and profits to e.g. newspaper industry if it manages to exploit the medium accordingly. Porter (2001), points out that the Internet is an extension - instead of a substitute - of conventional business. He further clearly states that the Internet is not disruptive to a firm's strategy but rather complementary, i.e. in the case of the newspaper industry it serves as an additional channel of news and information distribution. Porter argues furthermore that firms must transfer traditional ways of adding value to their product into the online channel to stay in line with their established strategy.

Our research question is to what extent the population of the Czech Republic is willing to pay for online content.

Methodology The aim of this thesis is to further explore the area of online content with relation to the willingness to pay for it from the consumers. With focus on the Czech Republic this thesis will work with data gathered through an online survey.

Outline

1. Introduction
2. Literature Review
3. Theoretical Concepts

4. Data and Empirical Model

5. Results

6. Conclusion

Core bibliography

1. HANEMANN, W. M. (1991): "Willingness to pay and willingness to accept: how much can they differ?" *The American Economic Review*: pp. 635–647.

2. KNETSCH, J. L., & SINDEN, J. A. (1984): "Willingness to pay and compensation demanded: Experimental evidence of an unexpected disparity in measures of value." *The Quarterly Journal of Economics*, **99(3)**: 507–521.

3. SHOGREN, J. F., SHIN, S. Y., HAYES, D. J., & KLIEBENSTEIN, J. B. (1994): "Resolving differences in willingness to pay and willingness to accept." *The American Economic Review*: 255–270.

4. WERTENBROCH, K., & SKIERA, B. (2002): "Measuring consumers' willingness to pay at the point of purchase." *Journal of Marketing Research*, **39(2)**: 228–241.

5. CHYI, H. I. (2005): "Willingness to pay for online news: An empirical study on the viability of the subscription model." *Journal of Media Economics*, **18(2)**: 131–142.

6. LOPES, A. B., & GALLETTA, D. F. (2006): "Consumer perceptions and willingness to pay for intrinsically motivated online content." *Journal of Management Information Systems*, **23(2)**: 203–231.

7. BAKKER, P. (2008): "The Simultaneous Rise and Fall of Free and Paid Newspapers in Europe." *Journalism Practice*, **2(3)**: 427–443.

8. BAKKER, P. (2002): "Free Daily Newspapers – Business Models and Strategies." *The International Journal on Media Management*, **4(3)**: 180–187.

9. CHYI, H. I. & LASORSA, D. L. (2002): "An Explorative Study on the Market Relation Between Online and Print Newspapers." *The Journal of Media Economics*, **15(2)**: 91–106.

10. DOU, W. (2004): "Will Internet Users Pay for Online Content?" *Journal of Advertising Research*, **44(4)**: 349–359.

11. WANG L. C., ZHANG, Y., YE, L.R. & NGUYEN, D.-D. (2005): "Subscription to Fee-Based Online Services: What Makes Consumers Pay For Online Content?" *Journal of Electronic Commerce Research*, **6(4)**: 304–311

12. PORTER, M. E. (2001): "Strategy and the Internet." *Harvard Business Review*, **79(3)**: 62–78

_____   _____
              Author   Supervisor

Chapter 1

Introduction

The Internet usage and accessibility went through an outburst in recent decades, reaching 78 percent of Internet users in developed and 32 percent in developing countries in 2014 and still growing (ITU 2014). The virtual reality allows people to access any information they want at any time they desire. This information, on one hand, makes their lives a lot easier—information is widely available, consumers can order services, compare prices, or find substitutes for goods they desire without much effort. On the other hand, companies have to face the toughening competition, rivalry and low barriers to entry in the industry of Internet services, putting enormous pressures on firms to take part in a destructive price competing race (Porter 2001). One of such industries is the newspaper publishing industry that with its traditional channel of publishing newspaper printed on paper is facing the widening competition of readily available online news. The online media is mostly free of charge, because even though the online usage increases, it proved to be difficult to monetize online users (Kohut *et al.* 2010). Therefore, companies need to properly adjust their strategies to stay competitive and deploy the Internet technology accordingly (Porter 2001). Studies show that for traditional companies, the Internet is not a disruptive force that would only take consumers away to the new product, but it is more of a complement to their traditional ways of competing and will not replace conventional ways of doing business nor overturn traditional advantages (Porter 2001; Dutta-Bergman 2004). Yet it seems to be more and more difficult to generate revenue as the paid newspaper usage is on the decline for many reasons, one of which is because of the consensus that "it is impossible to charge for general news content, because it is freely available in a similar form elsewhere on the Internet" (Herbert & Thurman 2007, pg. 215). Nevertheless,

the online publishers have found some audience that is willing to pay for their services.

This research paper attempts to explain key drivers for consumers' willingness to pay for online newspaper content in the Czech Republic. Empirical data was collected in the Czech Republic via an online questionnaire to analyse in local conditions the **Research question: Are inhabitants in the Czech Republic willing to pay for online content?** The newspaper industry was chosen consciously, on one hand, to narrow down the objective of the study, and on the other, to explore factors in an area that is not yet fully understood.

The thesis is structured as follows: Chapter 2 provides readers with a literature background covering the content of past studies of the willingness to pay in general and also the connection to the newspaper industry. Chapter 3 covers the theoretical framework necessary for developing a model to test the empirical data. Hypotheses of this thesis are also outlined in that section. Chapter 4 explains the empirical part of the paper, starting with the data collection, descriptive analysis of the collected data and continuing with the regression analysis. The results, along with limitations of this thesis, are presented in the following section Chapter 5. The last section Chapter 6 summarizes the findings.

Chapter 2

Literature review

2.1 Newspaper industry

2.1.1 From print to digital

More than a half of the world's population read newspaper daily, 2.5 billion in print, 800 million online (WAN 2014). There has been a 22 percent increase in online news readers to 44 percent between 2004 and 2010 according to Kohut *et al.* (2010). Overall readership of news media across all platforms has never been higher, thanks to new attractive choices of print and digital news media (OPA 2013). The digital revolution have caused the decline in readership of printed newspapers, dropping by 2 percent globally between 2009 and 2012 (WAN 2014). This may be attributed to the ageing population along with the fact that younger people read less in general (Dimmick *et al.* 2004), but also are more attracted to the free digital media (Mindich 2005).

Figure 2.1: Sources of News content
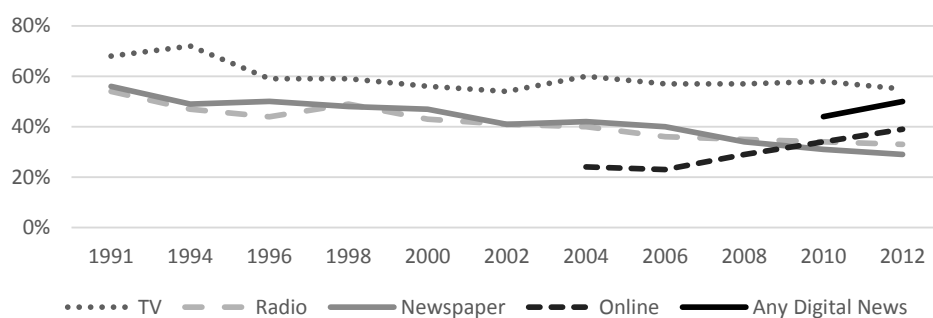
Source: PEW (2013)

One of the questions this paper tries to find an answer to is whether con-

sumers prefer accessing the information in form of printed media or in form of the digitally distributed online media. Previous research suggests that for newspapers that have both the printed and online version, readership bases overlap substantially (Chyi & Lasorsa 2002). Therefore, as Porter (2001) implies, the Internet is rather a complementary service to the traditional ways of doing business and not a disruptive force, allowing traditional newspapers to strengthen the relation with their readers by also offering digital news. This digital content can even attract additional readers, that would otherwise not read news at all (Picard 2001).

On the contrary, this complementarity and efforts to take an advantage of the Internet bring up a threat of loosing the readership base of the established printed newspapers and the connected sales and advertising revenues, if the online alternative with the same content and lesser, or even no price, is offered. The evidence against this argument was offered by Chyi & Lasorsa (2002), who have showed that the readership of the printed version of a newspaper is not reduced among readers of online news. Therefore printed counterparts still seem to remain competitive.

Chyi & Lasorsa (2002) further explore that users who read news online still prefer the printed format, given the same price and content. This suggests the perceived inferiority of online version of the newspaper and the online media in general, further studied and confirmed by (Chyi & Yang 2009). Moreover, Chyi & Lee (2012) in their study of stated preferences via telephone survey among Hong Kong citizens suggest that 58 percent of respondents still found the printed edition to be their favourite newspaper format in 2012. Nevertheless, even though users may still prefer the print, they seem to get along with digital media well, as illustrated in Figure 2.1.

2.1.2 From paid to free

The circulation of paid newspapers is on the decline in Europe (Bakker 2008). Picard (2001) argues that that free papers were able to attract a new reading audience that normally read newspapers only occasionally. Readers also seem to have changed their preferences. Bakker (2008) concludes that these preference changes can be attributed to the decrease in demand for specialized content, shifting the readership base to to the lighter news products in terms of coverage. However, without the specialized content, readers do not assign newspapers with a proper monetary value, considering it less valuable or

even not valuable at all, supporting the consensus of the Internet industry that
"content must be free, unless it is very specialized" (Carlson 2003, pg. 54).
Therefore the evidence suggests that readers do not just switch from printed
to online media, but do have a slightly different preferences when consuming
news.

The first free newspapers were introduced in Sweden in 1995 and since then
have found wide audience (Bakker 2008). Such newspapers have been growing
since and spreading to other European countries, yet after 2003 there has been
more newly published titles than new countries with free dailies (Bakker 2008).
This suggests that the competition in the newspaper industry is getting tougher
and companies have to face free substitutes to their traditional paid newspapers
either in form of printed or online newspaper titles. Free printed newspaper
similarly to online news portals have spread enormously in recent decades and
apparently "are here to stay" (Bakker 2002, pg. 186).

With the expansion of free newspapers, publishers face greater challenges
in how to generate revenue to run their business properly. Publishers therefore
consider various payment methods, as discussed in Subsection 2.1.3. Unfortu-
nately, running initially a free model and trying to switch to a paying model
later meets the unwillingness (Smith 2010) and a perceived unfairness due to
the established 'free mentality' of online content (Schonfeld 2001).

Switching to paying models also triggers the substitution effect (Chyi 2005),
since free newspaper content is plentiful online (Adams 2012). Setting up a pay-
wall causes readers to immediately move to reading other newspaper titles and
the demand for paid titles drops significantly (Hanemann 1991). The only way
that the substitution effect would not occur would be if all of newspaper pub-
lishers were to switch to a paying model at once, getting rid of free substitutes.
But since newspapers with free models can be profitable, it is unlikely that a
massive substitution from free models to paid models would occur.

2.1.3 Sources of revenue

The newspaper industry generates more than US$ 168 billion[1] of revenue annu-
ally (WAN 2014). Online publishers, on the other hand, are having difficulties
with monetizing the online content, as only about 7 percent of global newspaper
revenue come from digital sources (WAN 2014).

---

[1] 1 Czech Koruna equals approximately 0.039222 U.S. dollars as of April 2014 according
to the Czech National Bank.

Digital publishers with the 'noone would pay for it' consensus face the horizontal demand curve and the infinite elasticity of demand Chyi (2005). Once they try to increase the price, the substitution effect kicks in and the quantity demanded instantaneously drops to zero, because there are free substitutes. As mentioned previously, in the case of no free substitutes, people would be forced to use what they actually prefer and therefore decide, what are they willing to pay for.

Figure 2.2: Global Advertising Revenue
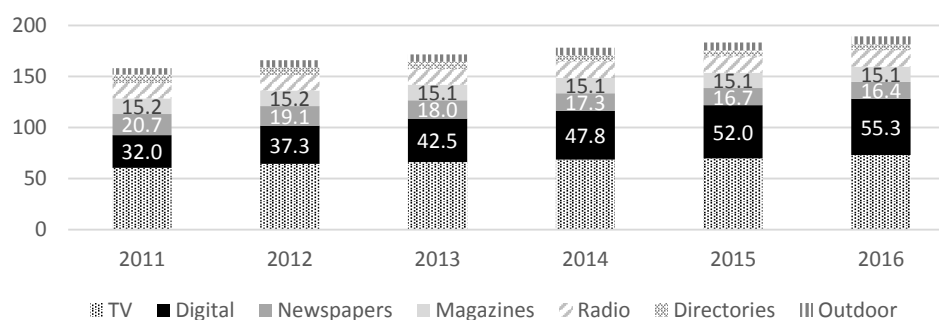
[US$ billion]
Source: PEW (2013), eMarketer."US Ad Spending Forecast." October 2012

The primary focus of a publisher is to generate an economic value, which "is nothing more than the gap between the price and the cost" (Porter 2001, pg. 65), generated by two major revenue streams—circulation (direct sales) and advertising revenues. On average, newspapers received 49% of their revenue from circulation and the rest from advertising in 2012 (OPA 2013).

Circulation and distribution

Newspapers are directly sold to consumers either via a subscription or single copy sales, creating the second largest revenue stream for publishers of the printed newspapers. Printed newspapers also have to be delivered via a distribution channel, which is closely connected to the sales process and is therefore discussed together in this section.

The delivery to subscribers is accounted for in the price of the subscription and is easily predicted. This is not the case for the free newspapers that have to rely on cheaper distribution channels, such as local public transport, shopping malls, universities and others (Bakker 2002). Picard (2001) argues that such distribution is very convenient, since people travelling or waiting in public places have the spare time to read such papers. Therefore the free printed

newspapers are able to find their audience, while taking the advantage of the cheap distribution.

Overall circulation revenues are declining (WAN 2014), but as Figure 2.3 illustrates, the digital circulation revenue is rising rapidly. The reason can be, that users may finally be prepared to pay for online content and to subscribe to online news services.

Figure 2.3: Global Digital Newspaper Circulation Revenue
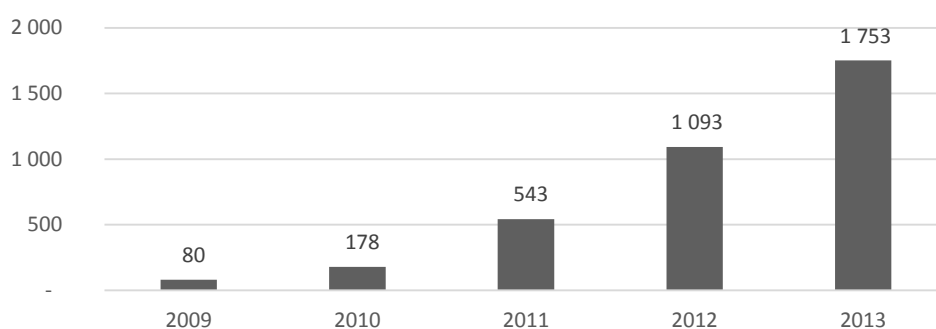
[US$ million]
Source: WAN (2014), PwC Global Entertainment & Media Outlook

Sales are still very important source of revenue for traditional printed newspapers. Yet with the emergence of free newspapers, the importance of the advertising revenue stream increased significantly. Moreover, as illustrated in Figure 2.3 from 2012 to 2014 the digital circulation revenue increased by 60 percent and makes approximately US$ 1.7 billion (WAN 2014).

Advertising

The advertising is the second largest revenue stream in the newspaper industry. For some—mostly free newspapers, it is the biggest revenue stream, as these newspapers do not generate revenue from sales (Bakker 2008). Research shows, that advertising revenues often do not meet the bottom-line that the newspaper company needs for its survival (Addison 2001; Dewan *et al.* 2002). Moreover, the circulation of printed newspapers is steadily declining, lowering the revenue from advertising in the print. According to WAN (2014), global newspaper advertising revenue in 2013 was US$ 84.79 billion, dropping by 13 percent when compared to US$ 97.44 billion in 2009.

In mid 2007, printed newspapers hoped to increase their revenue stream by taking an advantage of emerging Internet technologies. Publishers offered an online version of their newspapers, engaging in the online advertising business,

hoping to make up for losses from declines in circulation of printed newspapers. Unfortunately, it takes time to build an online brand and therefore the additional revenue from online advertising did not make up for the losses at the early stages of the expansion to the web (Farhi 2008). The transition to the digital content is rather slow as every US\$ 1 of digital revenue is accompanied by a US\$ 7 loss in print revenue (WAN 2014).

Advertising also repels viewers, further decreasing the overall revenue. Thus it may be even optimal to have a negative initial cash flows from having fewer advertising, yet being compensated for it later as a website gets more viewers that create future profits Dewan *et al.* (2002).

Figure 2.4: Global Digital Newspaper Advertising Revenue
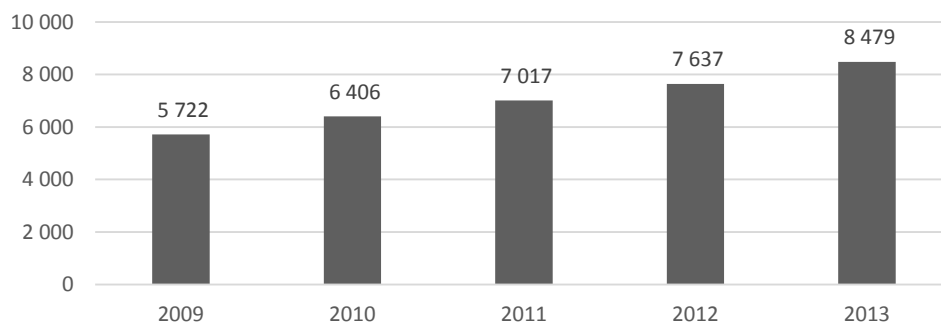
[US\$ million]
Source: WAN (2014), PwC Global Entertainment & Media Outlook

From 2012 to 2013 the global digital advertising revenue increased by 11 percent and made approximately US\$ 8.5 billion in 2013 (WAN 2014), as illustrated in Figure 2.4.

Payment models

In the newspaper industry, companies tried to impose alternative revenue generating models to boost their revenue stream. Mings & White (2000) specify these models such as a subscription model with regular payments to access the service, an advertising model, a transactional model, micro payments model or a bundled model. Strategies of companies differ as they try to convince readers to pay for digital access, but the subscription based model seems to be the most feasible model. In the US, the dominant paying strategy (39 percent share in paid models) is the metered subscription, that allows users to access specific amount of articles per given period (WAN 2014).

The subscription based model could make up for the losses due to declining

advertising revenues from printed media. The trend of turning a web based content to a subscription based content was observed in early years of 21st century by Wang *et al.* (2005) who explained that users were still not willing to pay for this model. The feasibility of the subscription model was further examined by Edmonds (2009) and Tom Rosenstiel (2012) that also concluded difficulties to generate revenue from this source for the digital media, even though it is the most frequently used model.

It remains unknown, why online news websites are unable to apply subscription models successfully even though 50 million of US readers still subscribe to printed newspapers (Farhi 2008). On the other hand, publishers do their best not to solely depend on the advertising revenue. According to the OPA (2013) Research, 95 percent of the leading digital brands have a subscription strategy. The research further suggests that introducing paid models as a part of the business strategy have overall positive effects on businesses. Initially the advertising revenue generated by visitors of a newspaper website goes down when the free access is limited, but the future revenue from subscriptions makes up for these loses (OPA 2013).

More recently, the digital newspaper industry seems to have stabilized in most European markets. Reuters (2014) reports that 11 percent of people have paid for digital news in the last year, either using one–time payment or have a subscription. BCG (2009) suggests that the maximum amount of money people are willing to pay for online news is very modest, around \$5[2] per month and only 2 percent of readers would be prepared to pay for unrestricted access to a website if a paywall was introduced.

Yet there is some evidence, that limiting the digital access can help publishers to boost their revenue, as was illustrated by the US news portal *Star Tribute*. Since they started to charge for the full access to the website, the readership faced a 12 percent decline, but subscription revenues exceeded the losses from decreased traffic by more than 15 times (Adams 2012).

Paid online newspaper models seem to be—slowly but continuously spreading and penetrating the newspaper industry. The readership base have not yet fully accepted this new environment, but it seems that now such models are more feasible than ever before. Publishers can eventually be able to monetize their readers, even though changing users' perception of online content takes time. Nevertheless, if websites do provide high quality products with a certain

---

[2]CZK 123 as of April 2015, according to the exchange rate of the Czech National Bank.

degree of uniqueness, that are both useful and interesting for the audience, it will eventually find the readership that will pay.

2.1.4 Newspapers in the Czech Republic

The descriptive statistics about the Internet population and the engagement in reading news in the Czech Republic comes from the research 'Media Projekt' conducted by the Association of Publishers[3] in the Czech Republic, therefore further statistics refer to this source 'MediaProjekt (2013)', unless stated otherwise. The data was collected during the third quarter of 2013 with 4 925 respondents with the access to the Internet, focusing on aspects such as usage, perceived benefits, or consumer preference for digital and printed news. The data refers to inhabitants of the Czech Republic.

The situation in the Czech Republic is consistent with global newspaper trends outlined in previous sections. The Figure 2.5 illustrates the total average circulation of printed dailies in the Czech Republic. Paid newspapers are declining, which is consistent with the points made in the Subsection 2.1.2. Apparently, the circulation of free printed dailies have stabilized, due to the fact, that for the past 5 years there have been exactly 2 of them—*E15* and *Metro*. Associated advertising revenues from circulation are declining as well, according to the data summarized in Table 2.1. Again, this fact is consistent with global trends discussed in Section 2.1.3.

Figure 2.5: Total average circulation of printed dailies in the Czech Republic
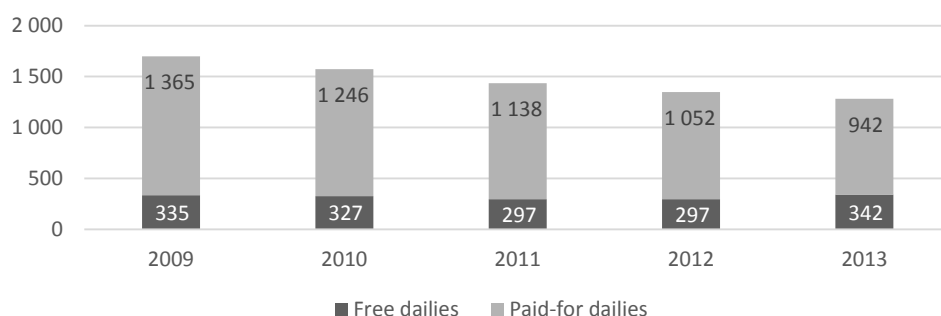
[Thousands]
Source: WAN (2014)

Today about 75 percent of inhabitants (between 12 and 79) have an access to the Internet, 74 percent know about the digital alternatives to classical

---

[3]Unie Vydavatelů, `http://www.unievydavatelu.cz/cs/home`.

printed newspapers, 39 percent have a reading device (phone, tablet, e-reader), 6 percent have ever read an article in electronic form, 6 percent pay for such content and just about 1 percent pay for online newspapers access. In Figure 2.6 we can observe the comparison to the rest of European countries in terms of paying for content online, showing the Czech Republic to be one of the least WTP countries. Quite the opposite seems to be northern-european countries that have the highest percentage of paying population, but also the reading population in general Bakker (2008).

Figure 2.6: Individuals using the Internet to buy or order online content in 2014 (age 16 to 74)
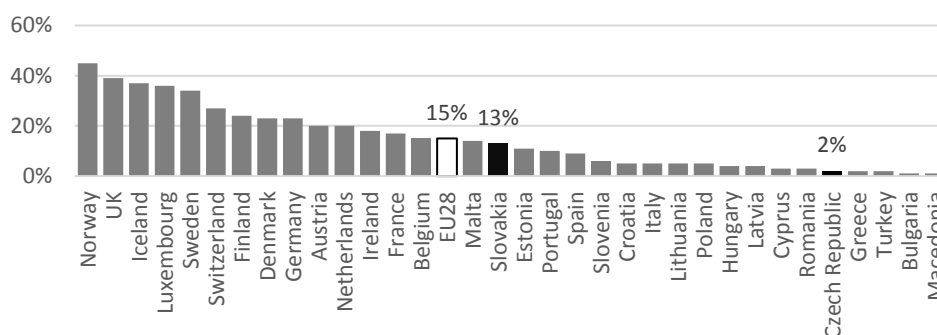
Source: Eurostat (2015)

In the Czech Republic, 56 percent of inhabitants still preferred the print in 2013. Another 33 percent prefer the print, yet plan to move to the digital media soon. Just about 3 percent read only the digital media, 7 percent do not read news at all. Readers mostly read the print because they are used to it. But since we have already established in Subsection 2.1.1 that the online content is supposed to cost less, the price competition can cause other readers to be attracted to the online media. In fact, about half of respondents who read the digital media are already attracted by the lower price.

About 30 percent of habitants use a smartphone, a tablet or an e-reader, which is almost twice as much as the year before that, suggesting an expansion of digital devices. Most readers that use digital devices are within categories 20-29 years, university educated, or living in Prague. In the Czech Republic, 61 percent prefer the web edition, 12 percent the pdf, 6 percent the application. 49 percent would like to access an extended version of their favourite newspaper (images, audio, video) online, 23 percent summarized information, 23 percent word-to-word copies and 6 percent mobile applications.

2/3 of internet users have visited some online news site in the previous 7

days, of which 94 percent is not WTP for online content at all, 79 percent is not WTP in the nearby future. If respondents were to pay, it would be for the content from their own expertise, suggesting WTP for specialized content.

Table 2.1: Czech Republic in numbers

Czech Republic data	2009	2010	2011	2012	2013
Internet users[1]	6 757	7 243	7 695	7 909	7 797
Total advertising expenditures[2]	1 072	1 130	1 160	1 160	1 149
Newspaper advertising expenditures[2]	153	130	118	113	112
Paid-for and free dailies[1,3]	1 700	1 543	1 435	1 349	1 284
Paid-for dailies[1,3]	1 365	1 246	1 138	1 052	942
Free dailies[1,3]	335	327	297	297	342
Non-dailies[1,3]	477	456	417	–	–

[1][Thousands], [2][US\$ million], [3]Total average circulation
Source: WAN (2014)

It seems that in the Czech Republic readers are strongly against paying for Internet content. Yet even though these descriptives provide us with an estimation of the real population to which we can compare our sample, they do not explain what drives people to pay. This thesis tries to find relationships among different variables that may be predictors of WTP.

2.2 Willingness to pay

So far this thesis has narrowed down the studied topic, and tried to explain, where the newspaper industry is heading and what challenges it faces. The established declining advertising revenue is to be compensated for by charging Internet users for what they consume. Yet the overall evidence suggests a rather disappointing execution of charging for the content. A better understanding of Internet users, and what drives their decision making when considering the actual purchase on the Internet has to be provided.

People tend to spend the highest amount of money online for entertainment purposes (Goyanes 2014), where the reasons seem straightforward. But what drives people to spend their money for information, or more precisely for the newspaper content? Since most of the online news portals provide information for free and this information tends to be similar across different publishers, the question arises, whether it is even possible to charge for general news content.

Consumers often take online services for granted and therefore when asked to pay, they perceive unfairness and are not willing to pay (Schonfeld 2001).

But does this translate into general unwillingness to pay or are there ways to monetize online newspaper content for companies? Unfortunately consumers seem to be irrational since according to Chyi & Lasorsa (2002), they are not actually willing to pay for what they use and also do not use what they prefer, shedding little light on the matter. This section will try to uncover the topic of consumers' WTP and identify key drivers that affect the paying intent, which is a basis for the chosen research method of this thesis.

2.2.1   Concept

Willingness to pay has been studied by many economists, psychologists and marketing researchers in attempts to estimate the demand for goods and find the optimal pricing strategy, as even minor changes in prices can have notable effects on overall profits and revenues (Marn *et al.* 2003). There are several existing research techniques to measure consumers' WTP, that differ in the way of acquiring the preferences and in whether they are being collected at the point of actual purchase or not (Wertenbroch & Skiera 2002).

Preferences can be classified as either *stated*, obtained by direct and indirect surveys, or *revealed*, collected as actual responses to the price in actual markets or via created experiments (field, laboratory, or auctions), as classified by Breidert *et al.* (2006). In practice, the collection of revealed preference data is costly and thus cost-efficient surveys dominate the scene of studying WTP. A quick overview of most commonly used techniques that attempt to analyse the collected data about consumers' WTP follows.

Considering various experimental auction models to reveal consumers' WTP *a Vickrey auction* arises among first. In this method bidders submit written bids without knowing bids of other people that participate in the auction. The highest bidder wins, but pays the price of the second highest bidder, the dominant strategy being to bid the true amount a participant is WTP (Vickrey 1961). There are some limitations to this type of technique, as Kagel & Levin (2002) explain, since bidders often strategically bid over their true WTP to increase their chances of winning, while ending up paying the lower price.

Another method is called *the Becker–DeGroot–Marschak (BDM)* method, where a subject formulates a bid that is then compared to a price generated by a random number generator. If the bid is greater than this random price, the

bidder receives the item for the generated price. If the price is lower, bidder pays nothing and receives nothing (Breidert *et al.* 2006). This method seems to be feasible in certain 'at the point-of-purchase' settings, as it incentivizes consumers to bid the true WTP value, while not suffering from the overbidding bias, as the Vickrey auction does (Breidert *et al.* 2006). Yet according to Wertenbroch & Skiera (2002) the BDM method produces lower WTP prices than under hypothesized survey questions, because participants actually pay for goods, therefore the method also have some limitations that have to be accounted for.

There is also *the reverse-pricing method*, in which buyers state their bidding price that is then compared to a threshold value stated by the seller and everyone bidding above this threshold value pays his actual bid and receives the good. Yet it is argued, that this model incentivize to lower buyers' WTP amount to get a better deal (Breidert *et al.* 2006).

Moving to the survey instrument and comparing direct and indirect survey methods in terms of reliability, Breidert *et al.* (2006) advises against the use of direct surveys for several further stated reasons. Asking the price directly causes an unnatural focus on the price and displaces the importance of the product itself. Customers also have no incentives to reveal their WTP and even if they reveal their true WTP, there is no guarantee that this translates into a purchasing behaviour. Asking a WTP value directly means a challenging task for respondents, as they often misjudge the value of the product, creating unstable estimates of WTP.

Indirect surveys provide respondents with a cognitively easier choices, as they are presented with some sort of rating or ranking for several different alternatives of a product, revealing their WTP indirectly. Following methods study these indirectly-collected stated preferences.

A contingent valuation requires respondents to state their WTP as the price they would pay, meaning an open-ended valuation. The other possibility is giving single or repeated choices at given prices, meaning a close-ended contingent valuation (Wertenbroch & Skiera 2002).

A conjoint analysis is another widely used practical tool designed to determine perceived differences in utility of consumers' WTP, as respondents are offered several alternatives that attempt to reveal their true preferences (Wertenbroch & Skiera 2002). There are several ways of doing a conjoint analysis, most commonly used being to ask respondents to choose among alternatives, or ask for an ordinal interval-scaled equally spaced measures of preferences, that are

further evaluated by OLS regression or ANOVA (Breidert *et al.* 2006). Yet even conjoint analyses have several draw-backs, one of which is not including the decision-making in survey processes.

Other limitations to these survey-based researches are hypothetical questions asked for contingent valuations and conjoint analyses, as they do not incetivize participants to reveal their true WTP and only ask for hypothesized preferences (Hoffman *et al.* 1993). Nevertheless, deeper insights into feasibility of these study methods in online environment need to be provided.

2.2.2 WTP for online content

Even thought there are many different methods to measure consumers' WTP that work well for products in a classical economic environment, the situation in a virtual market with online content differs from these traditional procedures, as there is little uniqueness in goods, there is a close-to-unlimited amount of substitutes accessible on demand, and the goods are perceived as inferior to other physical alternatives (Dou 2004; Chyi & Yang 2009).

Still, the newspaper industry tries to refinance its online production through pay-for-content models, meeting with consumers' weak paying intents for digital goods and online newspaper formats (Bleyen & Van Hove 2007; Chyi & Yang 2009; Chyi & Lee 2012; Adams 2012). The unwillingness to pay for online services is observed among various Internet services, as was well illustrated by the company Yahoo! that started to charge a fee for its online auction listing in 2001 and faced a 90 percent decrease in the volume of listings (Schiffman 2001).

Porter (2001) implies the importance of an established brand and the difficulty to build internet brands due to the lack of a physical presence and a direct contact. The rationality of consumers and how they formulate the actual amount they are willing to pay for online products still seems to be unexplored, especially for new products, where the pricing is difficult (Lopes & Galletta 2006). To further support Porter's point, Lopes & Galletta (2006) argue that since the Internet provides us with a competitive area with low barriers to entry, the reputation of a provider of the service plays a key role in getting consumers interested in provided services. Therefore even if the product is of a higher quality and consumers expect to pay a higher price (Zeithaml 1988), for new sites the situation is a bit different, as consumers will not assign a newly

created site of an exceptional quality with a proper monetary value because of
its lacking reputation (Lopes & Galletta 2006).

2.2.3   Previous studies of WTP

WTP has been subject to many studies in various fields. More recently with the
transition to digital technologies the studies focus more on the online content.

Hanemann (1991) and Shogren *et al.* (1994) tried to find the difference
between WTP and the Willingness to accept (WTA) and their relationship when
different types of goods are considered. Hanemann (1991) stressed out the
importance of substitution effects, that highly impact the convergence in WTP
and WTA values. Shogren *et al.* (1994) concluded that for market goods with
close substitutes (candy bar, coffee mug), the WTP and WTA converges, yet for
non-market goods with imperfect substitutes (reduced health risk), the WTP
and WTA diverges.

Schwer & Daneshvary (1995) studied WTP for Public Broadcasting Stations
(PBS) in a situation, when public funding would be reduced and the station
would be forced to charge their viewers, conceptually being similar to paying
for online news, funded by advertising. In their study Schwer & Daneshvary
(1995) asked a hypothetical question in a survey to reveal the WTP for public
broadcast of local residents. To evaluate their results Schwer & Daneshvary
(1995) used several *regression methods* (ordinary least squares, restricted least
squares, and restricted Tobit regression) and identified independent variables
that influence WTP as demographic factors, a preference for a substitute, and
a TV use.

Young *et al.* (1998) studied the discrepancy between a reported purchase
intention and an actual behaviour. They have developed a model for mea-
suring a binary response for WTP, such that 1 indicates consumers' intention
to purchase a product within a designated time horizon and 0 if not. Then
they compared it to the real behaviour data, again with 1 signalling a pur-
chase within the time horizon and 0 otherwise. For the case, that these two
are equal, *a logistic regression* and *a binary regression* with predictors such as
demographics (age, gender or job category) produce good results. Yet since
surveys usually only measure proxies for the actual WTP that do not perfectly
match the actual purchase intentions, the usage of *a corrupted binary regres-
sion model* or *a hierarchical Bayes estimation* (for measuring multiple brands
from a common product class) was suggested for these purposes.

Sillano & de Dios Ortúzar (2005) studied the increasingly popular *mixed-logit models* to measure stated WTP in a procedure of estimating individual-level parameters. They were able to obtain useful results even with an extremely small sample (75 individuals), suggesting the emphasis to be put on the collection of high quality data, rather than preoccupation with the sample size.

Chyi & Lasorsa (2002) conducted a survey to compare the preference for online and printed newspapers in 2002. They concluded that printed newspapers was preferred among readers, yet identified a substantial overlap of online and print readership for local dailies.

Chyi (2005) studied the viability of the subscription based model for online news. Data was collected via telephone survey among Hong Kong citizens and analysed by *a hierarchical multiple regression analysis*, showing showed a very modest paying intentions, driven by the age, internet use, and whether consumers prefer print or online media, while leaving income as insignificant.

The subscription based model for Web content was also studied by Wang *et al.* (2005) that collected a survey-based data and evaluated it using a *multivariate analysis of variance* instrument, concluding a perceived unfairness when consumers were asked to pay for digital content, but showing a positive relation between WTP and the perception of convenience, essentiality, added-value, service quality and usage-rate.

Lopes & Galletta (2006) tried to explain the increasing popularity of internet subscriptions as an alternative source of revenue for internet companies that offer intrinsically motivated content (sites about education, sports, movies, books, etc.). Data was collected via an online survey, further evaluated by *structural models* that identified expected benefits to be main drivers of consumers' WTP, while leaving the perceived quality and the provider reputation to have only an indirect impact.

Chung (2008) tried to find what interactive features, if any, of online Newspapers do consumers actually use. *A regression based analyses* showed a rather infrequent use of interactive features, suggesting to focus on building a brand instead of trying to offer all types of interactive features to engage readers.

Elvestad & Blekesaune (2008) evaluated data from 23 European countries on newspaper reading and using *a multilevel analysis* explained the variation in newspaper reading. This variation was attributed to age, gender, education level, household income, but also showed different effects of variables across different countries. Their study also identified some common trends in newspaper use—men spend more time reading than women, highly educated people read

more, and people living in high income households seem to spend more time on reading too. These findings are relevant, as this thesis attempts to find key drivers for a specific county, resulting in possible differences in newspaper use and WTP for other nations.

Kopits *et al.* (2011) used *a multivariate logistic regression* to determine factors that influence WTP for genetic testing for Alheimer's disease, using the demographic variables (age, gender, race, income, or education) and also preferences regarding knowing of future health status measured on likert scales.

Chyi & Chadha (2012) in a survey based experiment identified predictors of an electronic device ownership and multi-platform news consumption, showing that people already use electronic devices for accessing news content on a weekly basis. Therefore digital technologies seem to keep to penetrate the population successfully.

Chyi & Lee (2012) formed *a structural model* to interrelate the preference, the use and the paying intent for online news, with attempts to help identify why are online news companies having difficulties with monetizing their content. They concluded that consumers do not use what they prefer and are not willing to pay for what they use, showing a certain degree of irrationality of consumers when forming their paying intentions for online content.

According to Chyi (2012), "how are users charged does not make much difference—whether they are charged does" [pg. 227]. In their study they tried to identify predictors of the paying intent for multi-platform newspapers, based on a web survey data with likert question to measure paying intent, open-end question to find the amount respondents are WTP for different formats, and a several scenarios to evaluate different paying models. *A conjoint analysis* and *regression analyses* were used to compare the results. It seems, that multi-platform news are already a reality, yet users are not yet ready to pay for what they use.

Lin *et al.* (2013) tried to explain a 'free mentality' phenomenon of the Internet, according to which the online content should be free of charge. Based on an online survey (likert scaled questions), using *a regression analysis* they measured consumers' attitude towards paying for online music, concluding a negative impact of free mentality on attitude towards WTP. Even though the concept of accessing online music slightly differs the concept of accessing online news , the free mentality phenomenon is well applicable for any charged online content.

Goyanes (2014) further explored factors that influence WTP for online news,

based on the theory of the economics of information. In his study, he performed *a logistic regression analysis* on a telephone survey data. The model was constructed based on a dichotomous dependent variable—the paying intent, coded as 1 for respondents that have ever purchased an online newspaper. A relation was found between the dependant variable and demographics (age showed a negative relationship, income a positive), but also showed significance of previous purchases for other digital products, such as movies or TV content, games, music. Other than that, social network usage also seems to increase the probability of paying for online news. However, Goyanes found variables gender, and games and music purchases to be insignificant. The study concludes that Internet users are still much more likely to pay for entertainment content than for knowledge-based content such as online newspapers.

2.2.4 Approach of this thesis

Based on the previous studies summarized in Subsection 2.2.3, an online survey was created to collect stated-preference data about the WTP for online news content in Czech Republic for purposes of this thesis, as it is the most widely spread method that targets Internet users directly. For more information about the data collection, question choices and descriptive statistics, please see Chapter 4.

Considering previously used methods for evaluating paying intentions for online content, *a logit estimation* will be performed to identify what factors influence the probability of paying for online news content. More information about the hypotheses and theoretical concepts of the model, along with the comparison to other models considered, are stated in Chapter 3. The data collection, variables of interest, and the model description is then summarized in the following Chapter 4. In Chapter 5 the results of the analysis, along with a discussion and limitations, are provided.

Chapter 3

Theoretical Concepts

3.1 Paying intent

The paying intent for the online content is the key variable that corresponds directly to the WTP and therefore is to be studied by this thesis. The variable was identified by previous researches of the studied topic (Chyi 2012; Chyi & Lee 2012; Kopits *et al.* 2011; Wang *et al.* 2005; Sillano & de Dios Ortúzar 2005; Young *et al.* 1998; Schwer & Daneshvary 1995) and collected by asking the question "Some newspapers are considering to charge users for online content. Would you pay for the content on their website?". The responses were set up as a likert scale responses ranging from 1 (Strongly disagree) to 4 (Strongly agree) as the WTP concept does not allow for people in the middle that "do not care". With this framework, we assume that everyone has an opinion regarding where their money goes, which seems to be a rational assumption. Paying intent represents the dependent variable in our model representing consumers' WTP that is to be predicted. Yet due the necessity to make an opinion regarding WTP, the data was filtered for responses $1, 2$ to correspond to unwillingness (assigned value 0) and responses $3, 4$ to correspond to willingness (assigned value 1).

3.2 Limited dependent variables

The theoretical framework for the following section refers to Wooldridge (2012). Bold variables represent the vector notation, such that $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_n)$ and similarly for other variables in this section.

The concept of WTP divides the population on those with paying intentions

and those without any. Therefore the variable that we try to observe is limited in values it ranges from. A such type of dependent variable is referred to as a Limited Dependent Variable (LDV) and needs a special treatment, because we want to predict fitted values in a given range. There are several possibilities for LDV to range from—binary (values 0 or 1), count (1, 2, 3, . . .), truncated (observations above or below threshold values are excluded from the sample), censored (some variables cannot be observed under some conditions), or other types of variables. Since our dependent variable equals $intent = 1$ for consumers that are willing to pay and $intent = 0$ otherwise, we will focus on a binary response variable that leads us to following models.

Naturally, the simplest way to estimate a binary response variable is using the Linear Probability Model (LPM) that uses the usual Ordinary Least Squares (OLS) framework. In such case, the expected value of the dependent variable can be also interpreted as a probability of the event happening ($y = 1$), formally for observation i: $E(y_i|\boldsymbol{x_i}) = Prob(y_i = 1|\boldsymbol{x_i}) = (\beta_0 + \boldsymbol{\beta}\boldsymbol{x_i})$. The main problem with this model is that the fitted values can actually end up bellow 0 or above 1, which makes little sense when considering the WTP framework, interpreted as probability of paying for content. Moreover, under this model disturbances are heteroscedastic and all partial effects of independent variables are constant.

In order to limit the values of the dependent variable to range from $\langle 0, 1 \rangle$ we have to use a nonlinear function instead, given by Equation 3.1.

$$Prob(y = 1|\boldsymbol{x}) = G(\beta_0 + \boldsymbol{\beta}\boldsymbol{x}) \tag{3.1}$$

The desired function G has to fulfil given conditions and range strictly above 0 and below 1. There are two most common functions that satisfy these two conditions—the standard logistic and the standard normal random variables.

Logit model

The first one is the logistic function G with a CDF similar to the standard logistic random variable.

$$G(z) = \frac{exp(z)}{1 + exp(z)} \tag{3.2}$$

The function $G(z)$ satisfies desired conditions, for $z \to -\infty$, $G \to 0$ and as $z \to \infty$, $G \to 1$. By plugging Equation 3.2 into Equation 3.1 we obtain the

logistic regression model, referred to as *a logit model*, given by Equation 3.3.

$$P(y = 1|\boldsymbol{x}) = G(\beta_0 + \boldsymbol{\beta x}) = \frac{exp(\beta_0 + \boldsymbol{\beta x})}{1 + exp(\beta_0 + \boldsymbol{\beta x})} \tag{3.3}$$

It is used frequently due to its easy computation. Yet the interpretation of coefficients is not straightforward—to get a partial effect of x_j on the dependent variable we have take a partial derivative w.r.t.[1] x_j (see Equation 3.4).

$$\frac{\partial P(\boldsymbol{x})}{\partial x_j} = g(\beta_0 + \boldsymbol{\beta x})\beta_j, \qquad \text{where } g(z) = \frac{dG}{dz}(z) \tag{3.4}$$

Since $G(\cdot)$ is a strictly increasing function, $g(z) > 0$ for each z. The sign of the partial effect is therefore given by the sign of β_j directly. This general derivation of marginal effects of dependent variables is also well applicable for the probit model.

Probit model

The second model takes for the function G a standard normal CDF.

$$G(z) = \Phi(z) = \int_{-\infty}^{z} \phi(v)dv, \tag{3.5}$$

where $\phi(z)$ stands for the standard normal density: $\phi(z) = \frac{1}{-2\pi}exp(\frac{-z^2}{2})$. Combining Equation 3.5 with Equation 3.1 we obtain *the probit model*, given by Equation 3.6.

$$P(y = 1|\boldsymbol{x}) = \Phi(\beta_0 + \boldsymbol{\beta x}) \tag{3.6}$$

Maximum Likelihood Estimation

Due to the non-linear nature of probit and logit models, we have to estimate them using the Maximum Likelihood Estimation (MLE). For a sample of size n we need the density of y_i given x_i, given by Equation 3.7.

$$f(y|\boldsymbol{x_i}; \boldsymbol{\beta}) = [G(\boldsymbol{x_i\beta})]^y[1 - G(\boldsymbol{x_i\beta})]^{1-y}, \qquad y = 0, 1, \tag{3.7}$$

---

[1]Refers to a common abbreviation of 'with respect to'.

the log-likelihood function for all observations being represented by Equation 3.8.

$$\mathcal{L}(\boldsymbol{\beta}) = \sum_{i=1}^{n} y_i log[G(\boldsymbol{x_i}\boldsymbol{\beta})] + (1 - y_i)log[1 - G(\boldsymbol{x_i}\boldsymbol{\beta})] \qquad (3.8)$$

The MLE of $\boldsymbol{\beta}$ for the sample of size n maximizes the log-likelihood given by Equation 3.8. The estimated coefficients of $\boldsymbol{\beta}$ are denoted by $\hat{\boldsymbol{\beta}}$. If the $G(\cdot)$ is the standard logistic CDF then $\hat{\boldsymbol{\beta}}$ represent *the logit estimator* and for the case of $G(\cdot)$ being the standard normal CDF, $\hat{\boldsymbol{\beta}}$ represent *the probit estimator*. Logistic slopes are interpreted as the effect of a unit change in X variable on the predicted log of odds, while other variables are constant.

Properties of probit and logit

Even though both models are rather similar, logit was traditionally used more often. Yet for today's application in economics the standard normal distribution of probit is more realistic.

Figure 3.1: Probit and logit CDFs
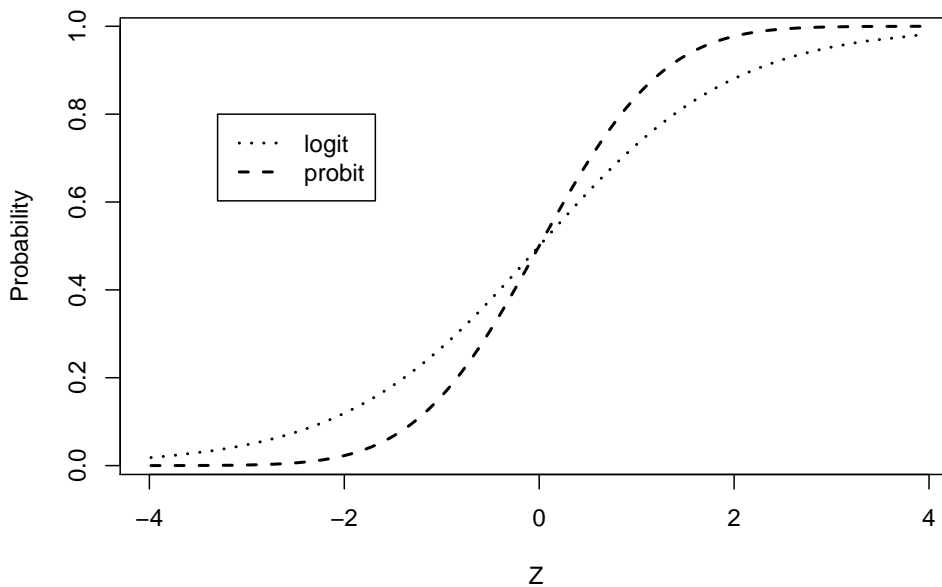
Source: Author's computations.

Logit and probit models can also be derived from a latent variable model.

$$y^* = \beta_0 + \boldsymbol{\beta}\boldsymbol{x} + e, \qquad y = 1 \text{ for } y^* > 0, y = 0 \text{ otherwise},$$

where y^* is the latent variable and y is the indicator function that takes the

value of 1 under the condition $y^* > 0$. Further the probability is derived as:

$$P(y = 1|\boldsymbol{x}) = P(y^* > 0|\boldsymbol{x}) = P[e > -(\beta_0 + \boldsymbol{\beta x})]$$
$$= 1 - G[-(\beta_0 + \boldsymbol{\beta x})] = G(\beta_0 + \boldsymbol{\beta x}),$$

which is the same as Equation 3.1. We assume that e is independent of $\boldsymbol{x}$ and that e has either the standard logistic distribution or the standard normal distribution. The error term e is symmetrically distributed around 0 in either case, which means that $1 - G(-z) = G(z)$ for all real z.

The theory of conditional MLE implies that the MLE estimates $\hat{\boldsymbol{\beta}}$ are consistent, asymptotically normal and asymptotically efficient. This allows us to easily derive asymptotic standard errors for estimates and therefore test single hypotheses using the asymptotic t tests. Therefore these models are suitable for our study and since probit and logit are rather similar, the more traditionally used approach—the logit estimation, is chosen, along with evaluating the limitations of the LPM.

Moreover, to test hypotheses about one or more coefficients we can use likelihood ratio of the likelihood in the model restricted by our hypothesis over the likelihood of the unrestricted model.

3.3   Hypotheses

Previous research concluded several various results. This thesis aims to further attribute to the evidence of what factors cause consumers to pay for online news content. Several key variables and hypotheses were outlined in accordance to the relevant literature.

3.3.1   Relevant factors

Devices with Internet access

With the rising usage of modern technologies, there are no longer just two ways to access information—the web and the print. Other methods, such as reading on a tablet device, a smartphone, a classical phone, a TV, or even a game console seem to have a success among readers (BCG 2009). Still, web is considered to be the least desirable format Chyi & Lasorsa (2002).

Publishers are therefore not limited just to the web when trying to take a strategic advantage. Mobile phones, tables and e-readers are being used more

than ever before. According to DCN (2014) 68 percent of smartphone users occasionally follow breaking news, and about 33 percent do this frequently. Consumers might be willing to pay for the mobile edition of the news, despite the unwillingness to pay for web edition.

Moreover, since most of newspapers are published across all platforms—mobile devices, the web or the print, the consumer can choose among any of these formats throughout the day. In Subsection 2.1.1 we have established, that consumers still prefer the print and are willing to pay for it. But what if we allow people to access the same information from their mobile phone, with the convenience of not having to physically go and get a copy of newspapers, would consumers be willing to pay? The following hypotheses is stated.

Hypothesis 1.a: *An ownership of an electronic reading device with Internet access positively affects WTP for online news.*

In Subsection 2.1.1 we discussed the preference for the printed edition of newspapers over online edition. Statistical evidence suggests similar finding in the Czech Republic, as explained in Subsection 2.1.4. But with the rapid emergence of the digital device usage, it is possible that the format preference started to shift towards the digital media. Still, those who consume the digital media will intuitively be much more WTP for online content than those who read the print, suggesting the following hypotheses.

Hypothesis 1.b: *Accessing the online news content positively influences consumers' WTP.*

Hypothesis 1.c: *Accessing the printed news content influences consumers' WTP.*

OPA (2013) Research views the mobile device ownership as essential to whether a digital subscription model is feasible. On one hand, tablets are seen as substitutes to reading a printed copy of a magazine or newspaper. On the other, smartphones face greater challenges, because they offer less screen reading space, but provide publishers with opportunities to differentiate themselves. The potential market is huge as the amount of people having a smartphone and actively looking up information via Internet increases. As Chyi & Chadha (2012) mention, multi-platform news may already be a reality as most of consumers already follow news on one of the electronic devices on a weekly basis.

Internet and news use

The amount of time consumers spend online is important by intuition, since we are discussing solely the virtual market with newspapers. The Internet use variable was identified and used in previous studies (Chyi 2005; Lin *et al.* 2013), implying its importance for the studied topic. Yet this direct relationship should be verified.

Since this thesis narrowed down the studied topic to WTP for online newspaper content, the engagement of the readers in reading is a relevant factor (Chyi 2012). Intuitively the more engaged people are in reading news, the more WTP they are. Following the arguments for the Internet use variable, following hypotheses are outlined.

Hypothesis 2.a: *There is a significant relationship between WTP and the time spend online.*

Hypothesis 2.b: *There is a significant relationship between WTP and the time spend on reading news.*

Previous engagement in online payments

Goyanes (2014) showed the significance of previous purchases for digital products, such as movies, TV content, games, or music to directly impact the WTP for online content. These consumers are comfortable enough to share their credit card information with providers of Internet services. They also do not feel inconvenience when paying for other online products, as they are already familiar with the process. This may translate to the WTP for the online news content directly.

Hypothesis 3.a: *Previous online payments have a positive influence on consumers' WTP for online news content.*

Yet to also focus on news consumption habits, whether readers are already paying for any type of newspaper is relevant.

Hypothesis 3.b: *Previous purchases of newspapers have a positive influence on consumers' WTP for online news content.*

Newspaper reader type and interest in news

There is a difference between what users consume most often and what are they read, because of their interest in the topic. Moreover, consumers do not always use what they prefer and are not WTP for what they use (Chyi & Lasorsa 2002). Furthermore, according to Lopes & Galletta (2006), the willingness to pay is not associated with consumers' interest. The exact opposite—strong influence of interest on use of both online and printed news was assigned by Chyi & Yang (2009). The intuition favours the latter—the more interested one is in newspapers, the more likely will he or she actually read and possibly pay for the content. Nevertheless, further evidence needs to be provided.

Hypothesis 4.a: *The field of interest have an influence on WTP for online content.*

Readers are most WTP for specialized online content according to various studies (Carlson 2003; MediaProjekt 2013; WAN 2014). Therefore what news content Internet users consume most often is a relevant factor influencing their WTP. This statement is to be tested by the following hypothesis.

Hypothesis 4.b: *The type of content consumers read most often have an influence on WTP for online content.*

3.3.2 Properties of paid news content

Measuring WTP is a complex task that should also evaluate what are people WTP for. The following factors were outlined based on previous studies and surveys.

Limited advertising

In Section 2.1.3 we identified that the amount of advertising on a website repels its users. Therefore consumers may be WTP for limiting these adds. Conversely, some customers may enjoy consuming certain local adds, since they may help them access local services. Both the supporters and opponents are reflected in our research.

Hypothesis 5.a: *A level of advertising does affect WTP for online content.*

Customized content

Additionally, newspaper readers may be WTP for content with a certain degree of a personal customization. In such case, users would be paying for an actual service, an added-value to the generally available content. A customized content also allows companies to collect more information about users of their sites and better focus different adds for different users, opening the possibility for additional revenues from advertising.

Hypothesis 5.b: *A customization of online content does affect WTP for online content.*

Good quality and detailed information

Since most people would pay for the specialized content that is related to the field of their expertise (MediaProjekt 2013), they may be WTP for the higher quality content. Such content provides users with an added-value when compared to other widely available news, allowing users to assign this content with the proper monetary value.

Hypothesis 5.c: *A higher quality content does affect WTP for online content.*

3.3.3 Demographic factors

Males read more

Studies suggest that sex plays an important role behind the reasoning, whether online users are likely to consume online media and furthermore pay for it. Poindexter *et al.* (2010) point out that males are more likely to be online readers. Chung (2008) suggests that males use more interactive features of websites, therefore might be more willing to start using new models and actually pay for content. According to Chyi (2012), males are more willing to pay for online content. Apparently sex is a variable of interest when talking about WTP. But whether these expectations apply to online readers in the Czech Republic should be tested, as according to Elvestad & Blekesaune (2008) factors that determine the newspaper use differ across countries.

Hypothesis 6.a: *Men are more WTP for news content.*

Younger audience

Younger readers tend to read less than they predecessors did (Dimmick *et al.* 2004). Moreover, younger people are less likely to engage in any type of news, regardless of the format (Mindich 2005; Prior 2007). They are also more attracted to free newspapers (Bakker 2007; Lauf 2001), further supporting the trends in the newspaper industry. The younger audience is also more likely to pay for the online news access (Goyanes 2014). In the newspaper industry digital subscribers are on average younger than those who subscribe to the print OPA (2013). Moreover, these subscribers are often new subscribers that would otherwise not be interested in reading the print, nor subscribe to the print. Therefore another hypothesis should be tested.

Hypothesis 6.b: *Younger readers are more WTP for online news content.*

Education

Statical data (BCG 2009; MediaProjekt 2013; WAN 2014) shows, that university educated users are most likely to pay for online content. This expectation is to be tested in the environment of the Czech Republic.

Hypothesis 6.c: *University educated Internet users are more likely to pay for online content than those with lower level of education.*

Occupation

Not only the acquired level of education plays a key role when deciding on factors influencing WTP, but also their occupation. In the Czech Republic, the demographic category most WTP are people between 20–29, being mostly students of universities (MediaProjekt 2013). This statement will be tested by the next hypothesis.

Hypothesis 6.d: *Students are most likely to pay for online content than employed readers.*

Income

Chyi (2005) suggests that income in not the key variable driving consumers' WTP in a subscription based newspaper model. On the contrary, Goyanes (2014) based on a telephone survey concluded that the income is a significant variable impacting the paying intent for online content. This discrepancy ought

to be tested and compared to the intuition that with increasing income, the amount of money one can spent for the newspaper consumption increases.

Hypothesis 6.e: *Consumers' WTP does depend on their income.*

3.3.4 Is the print preferred to the digital?

Following the reasoning from Subsection 2.1.3 that in the case of no free substitutes and introducing paywalls on news sites, people would be forced to decide what content they prefer to consume. This hypothesis was previously explored by Schwer & Daneshvary (1995) that measured WTP for TV broadcasting and Chyi (2012) who measured the amount of money users would pay for three different news platforms—the Web, the print and mobile apps. Yet since essentially, buying the access to either—the application or the website, allows users to access both of these platforms, only the difference between digital and print platforms is to be tested in this thesis.

Hypothesis 7: *If there were no free substitutes, Internet users would pay for digital access to news content less than for the printed counterparts.*

On one hand, Hypothesis 7 seems to be counterintuitive for Internet users that would rather pay for news in digital form, as they use Internet regularly. Yet accounting for the perceived inferiority of the digital content discussed in Subsection 2.2.2, these users can actually be WTP less for the digital edition than for the printed edition.

Chapter 4

Data and Empirical Model

4.1 Data collection

To examine proposed hypotheses an online survey was created to collect data and distributed among friends, family, and others (distant friends, work colleagues of friends and family, friends of friends, etc.), mostly using the Facebook social network or directly through e-mails. The survey was conducted during the spring of 2015, took approximately 5 minutes to complete and generated 263 responses. The primary language of the survey was Czech, since this thesis focuses demographic group of inhabitants of the Czech Republic. The full questionnaire is available in the Appendix A, having each of the questions also presented in its English translation.

Online surveys are becoming more and more popular due to theirs low costs, no interviewer bias, speed, or convenience to responders. Surveys of this type are also less intrusive for respondents, as they can choose when and where to actually complete the questionnaire, yet having the overall lower response rate, when compared to e.g. a telephone survey (Fisher 2005). Even though online surveys only collect data from online users, this study focuses on this group of respondents. The questionnaire is also answered only by people who are interested in the topic, therefore a certain degree of self-selection bias may be present in the sample.

4.2 Variables and measurement

Variables were selected based on previous studies that examined the willingness to pay of online users (Dou 2004; Chyi 2005; Chyi & Lee 2012; Lin *et al.* 2013;

Goyanes 2014) and further applied to the Czech Republic.

Dependent variable

Dependent variable *paying intent* was already identified and discussed in Section 3.1.

Independent variables

Following the arguments from Section 3.3, further stated independent variables were included in the model.

Device was coded as a binary variable (1 for "yes", 0 otherwise) and reflects whether consumers are in possession of a smartphone, tablet or e-reader with the Internet access. Such consumers are able to access the Internet content on demand which arguably have an influence on the WTP. *Access* variable measures the type of place from which consumers access the Internet most often (1 for "From a personal computer", 2 for "From a mobile device", 3 "From work/school" and 4 "From elsewhere").

OnlineNewsUse and *PintNewsUse* are the two variables that reflect the reading habits of consumers, as they indicate whether they accessed news in print or in digital edition in the past 3 days. These variables were again coded as binary variables (1 for "Yes I have accessed online/printed news in the past 3 days" and 0 otherwise).

InternetUse and *NewsUse* reflect the amount of time consumers spend on using the Internet or reading the news (in any format) per day. It is argued, that being more engaged in the virtual reality and spending more time online may impact the paying intentions for content, that is available exclusively in such an environment. Also, being more engaged in reading news content may have an affect on WTP. These variables were coded as continuous variables.

PastPayment, coded as a binary variable (1 for "yes", 0 otherwise) reflects, whether consumers have ever paid via the Internet. Arguably, those who do trust the Internet enough will engage in additional Internet purchasing activities. To further specify paying habits for newspapers *PastPaymentNewspapers* was also included among dependent variables that reflects the engagement in buying newspapers. The variable was coded as categorical to allow us to identify occasional buyers of newspapers (4-10 times per 3 months), regular buyers (10+ times) and those who buy newspapers rarely (1-3 times) or never (0 times).

Interest was included to get a better perception about the type of the consumer, or more precisely what is he or she interested in. Nevertheless, consumers often read more the content outside of their zone of interest, therefore variable *Reader* was also included. Both of these variables were coded as categorical and were collected by asking to specify "what are they most interested in" and "what do they read most often". There were 5 possible results, coded as 1 for "News content", 2 for "Tabloid content", 3 for "Sports content", 4 for "Specialized content", and 0 for "No content". These two variable were merged into a single variable *Content*, representing the interest variable with responses 0 for "Other content", 1 for "News content" and 2 for "Specialized content". For more information please see Section 4.4.

What consumers expect from paid content was also included among predictors. *LessAdvertising*, *CustomContent*, and *QualityContent* are all self explanatory. They were coded as categorical variables ranging from 1 to 5, 1 reflecting no perceived importance of this expectation and 5 a huge importance. We assumed that the distance across all points on the scale is the same and therefore treat the variables as continuous.

Demographic factors were also included. *Age* being a continuous variable measured in years, *Gender* a binary variable with 1 for a "male" and 0 otherwise, *Education* was expressed categorically matching the amount of years at school with the achieved level of education, such that for the Czech Republic "the elementary education" takes 9, "the high school education" 13, and "the university education" 18 years. *Occupation* was also included to allow to differentiate from students (corresponding category code 1), employed (category 2), self-employed (category 3) and unemployed (category 4). Another categorical variable included was *Income*, measuring the after-tax income of an individual in categories 0-10 (corresponding category code 1), 10-20 (category 2), 20-30 (category 3), 30-50 (category 4) and 50+ (category 5) in thousands of Czech Korunas. Other than that, *Nationality* was also included among predictors to ensure the Czech nationality of respondents, being coded as a categorical variable with possible responses Czech (category code 1), Slovak (category code 2), or Other (category code 3).

Other variables

Separated from independent variables, another two variables—*PrintSubscription*, *OnlineSubscription* were collected under hypothesized conditions that no news

content is free anymore. Under such conditions, questions were asked to assign a monetary value for a monthly subscription for a print and for an online edition of a favourite newspapers. The online edition included monthly access to both the website and application. Variables were coded as continuous amounts in Czech Korunas.

4.3 Empirical model

The model used in this thesis is based on a binomial logistic regression and analyses the probability of having paying intentions for online news content, as was outlined in Section 3.2 and Section 3.1. This logistic regression tests the probability of dichotomous events happening, in this case having the intentions to pay. By combining the theoretical framework from Chapter 3 with outlined independent variables from Section 4.2, the following model is to be predicted.

$$
\begin{aligned}
P(Intent =1|\boldsymbol{x}) = G(&\beta_0 + \beta_1 Age + \beta_2 Male + \boldsymbol{\beta_3} Educ + \boldsymbol{\beta_4} Occupation \\
&+\beta_5 Income + \beta_6 Nationality + \beta_7 Device + \boldsymbol{\beta_8} Access \\
&+\beta_9 OnlineNewsUse + \beta_{10} PrintNewsUse + \beta_{11} NewsUse \\
&+\beta_{12} InternetUse + \beta_{13} PastPayment \\
&+\boldsymbol{\beta_{14}} PastNewspapersPayment + \boldsymbol{\beta_{15}} Interest \\
&+\boldsymbol{\beta_{16}} Reader + \beta_{17} LessAdvertising + \beta_{18} CustomContent \\
&+\beta_{19} QualityContent)
\end{aligned}
\tag{4.1}
$$

The function G refers to the standard logistic function $G(z) = \frac{exp(z)}{1+exp(z)}$ for logit estimation and standard normal CDF $G(z) = \Phi(z)$ for probit estimation, as was explained in Section 3.2. Note that variables with bold coefficients are treated as categorical, therefore creating several dummy variables to cover effects of all possible outcomes. This step is provided by the software used for the computation.

The model is computed using the *Stata 12.0* software.

4.4 Data descriptive analysis

The total of 263 respondents completed the survey. There were no incomplete data, as all of the questions were compulsory. Table 4.1 summarizes the demo-

graphic characteristics of the collected sample. About 53 percent were male and 47 percent female. The average age of the respondents is just under 27 years. The overall age of respondents ranges from 15 to 61 years, indicating a large variety in the sample. On the other hand, almost 79 percent of respondents were 29 years old or younger. This causes a certain skewness in the frequency distribution with respect to age, due to mainly addressing peers—being students and their acquaintances. Similar trend can be seen in education levels, mostly representing high-school educated respondents that study college, as students represent almost 68 percent in the sample. This also corresponds to the income level, as about 57 percent have less than CZK 10000 per month. There are about 90 percent of Czech respondents in the sample, the second largest share was represented Slovaks with about 8 percent, still not significantly large group to be evaluated separately and compared with a sample consisting of just Czech respondents. Additional adjustments were made to the income category, merging groups 30000-50000 and 50+ into just one, due to the fact that only 4 respondents had an after-tax income of more than CZK 50000.

On average, respondents spend a little over 4 hours using the Internet per day and about ¾ of an hour reading news in either online of printed form. About 80 percent own a tablet, a smartphone, or an e-reader with an access to the Internet. About 70 percent connect to the Internet mostly from home using a PC, additional 28.5 percent use connect 'on the go' using a mobile device (tablet, smartphone, or e-reader) and the remaining 9.5 percent connect from work or school. As there was not a single response in the 'other' option, the variable was dropped. Reading online news is much more popular among respondents as about 87 percent have read an online newspaper in the past 3 days and only 24.7 percent have read a printed newspaper. In contrast, 66.5 percent of respondents have not paid for news content in the past 3 months, 22 percent have bought newspaper occasionally (1-3 times) and just above 33 percent have bought newspapers 4-10 or more than 10 times. It seems that people trust online payments and are already used to them as more than 82 percent of respondents have used the Internet to pay in the previous 3 months.

The majority of respondents do expect none or at least reduced amount of advertising, if they were to pay, with a mean value 4.11[1]. Even higher expectations are regarding the quality of the content with a mean value 4.47. An option to customize the content is considered to be least expected with a mean value 3.72.

---

[1]On the likert scale from 1 (Strongly disagree) to 5 (Strongly agree).

Table 4.1: Sample description

Variable	Frequency	Percentage
Gender		
Male	140	53.23
Female	123	46.77
Age		
15-29	207	78.71
30-49	43	16.35
50+	13	4.94
Income		
0 - 10 000	150	57.03
10 001 - 20 000	48	18.25
20 001 - 30 000	42	15.97
30 001+	23	8.75
Education		
Elementary	16	6.08
High-school	162	61.60
University	85	32.32
Occupation		
Student	173	65.78
Employed	83	31.56
Self-employed	5	1.90
Unemployed	2	0.76
Nationality		
Czech	234	88.97
Other	29	11.03
n	263	

Questions about the news interest and actual usage proved to be difficult to evaluate, not supporting the argument that consumers do not prefer what they use suggested by Chyi & Lee (2012). Due to the fact that most respondents answered the question "What content are you most interested in" the same way as "What content do you read most often", the latter variable was dropped. Moreover, answers consisted mostly of *news* and *specialized content*, therefore categories *tabloid* and *sport* were merged into *other* category of the newly created variable *Content*, representing the reading interest of respondents.

Chapter 5

Results

5.1 Logist regression results

In the theoretical part of this thesis, different models for estimation were outlined, along with the differences between LPM, probit and logit models. Since logit and probit yield similar results, author has decided to just report logit estimates. Comparison of several models can be further found in the Appendix B.

To get the best fitting model, maximizing the estimated likelihood, several models with various variables were performed and tested, using the likelihood-ratio test after each estimation, comparing the restricted and the unrestricted models, both being fitted by MLE. In total 8 models were tested and compared with other models. Results of all tests can be found in Table B.19 in the Appendix B. Using this method, Model (3) from the Table 5.1 was selected as the best fitted model.

It is important to note, that variables *Device*, *LessAdd*, *CustomContent* *QualityContent*, and categorical variables *Access* and *Content* proved not to significantly improve the estimated model. Therefore they were ruled out as not having an effect on the estimated dependent variable *Intent* that represents consumers' WTP.

$$
\begin{aligned}
P(Intent = 1|\boldsymbol{x}) = G(&\beta_0 + \beta_1 Age + \beta_2 Male + \boldsymbol{\beta_3} Educ + \boldsymbol{\beta_4} Occupation \\
&+ \beta_5 Income + \beta_6 Nationality + \beta_9 OnlineNewsUse \\
&+ \beta_{10} PrintNewsUse + \beta_{11} NewsUse + \beta_{12} InternetUse \\
&+ + \beta_{13} PastPayment + \boldsymbol{\beta_{14}} PastNewspapersPayment
\end{aligned}
\tag{5.1}
$$

Equation 5.1 summarizes the notation for the final model, after ruling out

Table 5.1: Logit regression, $\boldsymbol{\beta}$ coefficients

	(1) Intent	(2) Intent	**(3)** Intent	(4) Intent
Age	-0.0310	-0.0311	-0.0263	-0.0234
Male	0.944**	0.894**	1.283**	1.148*
Educ	0.0195	0.0275	0.0386	0.0167
Occupation				
Employed	-1.719*	-1.729*	-0.536	-0.554
Self-Employed	-0.612	-0.714	0.551	0.561
Unemployed	2.396	2.872$^+$	4.055*	4.433*
Income				
10001-20000	0.939*	0.941*	0.220	0.0510
20001-30000	1.711*	1.689*	0.535	0.166
30000+	2.933***	2.919**	1.643	1.374
Czech	-0.455	-0.459	-0.503	-0.291
Device		0.577		0.677
Access				
PC		-0.0890		-0.510
Phone/Tablet		0.403		1.022
OnlineNewsUse			0.656	0.821
PrintNewsUse			-0.196	-0.144
NewsUse			1.021**	1.059**
InternetUse			0.154$^+$	0.148
PastPayment			1.290$^+$	1.121
PastNewspaperPayments				
1-3			2.201***	2.485***
4-10			3.224***	3.435***
10+			2.468**	2.714**
Content				
News				-0.355
Specialized				-0.761
LessAdd				-0.0734
CustomContent				-0.113
QualityContent				0.203
_cons	-0.893	-1.465	-5.751**	-6.105*
N	263	263	263	263
Df	10	13	18	26
Chi-square	33.40***	35.26***	109.27***	116.02***
pseudo R^2	0.110	0.116	0.361	0.383

$^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

the redundant variables. This model (see (3) in Table 5.1) with 263 observations and 18 degrees if freedom is statistically significant with $\chi^2 = 109.27$, $Prob > \chi^2 = 0.0000$, and $pseudo$-$R^2 = 0.361$. Note that this is not the same R^2 that is used in the OLS estimation and does NOT imply the proportion of variance of the response variable explained by the predictors. In this case, we have the McFadden's R^2 that is computed such that $R^2 = 1 - \frac{ln[\hat{L}(M_{Full})]}{ln[\hat{L}(M_{Intercept})]}$, where M_{Full} is the model with predictors (or alternatively called the *unrestricted model*), $M_{Intercept}$ is the model without predictors (or *restricted model*) and $\hat{L}$ represents the estimated likelihood. Therefore when comparing the two models, a small ratio of log likelihood indicates that the full model is better than the intercept model. Therefore higher the McFadden's R^2, greater the likelihood of the model.

5.2   Hypotheses testing

There are three standard ways to test exclusion restrictions for coefficients of the logit model—a Wald test, an LR test, a Lagrange multiplier or a score test (Wooldridge 2012).

The Wald test is the easiest one as it only requires to estimate the unrestricted model (full model with all independent variables). The Wald statistic is a function of the difference in the MLE and the hypothesized value such that $W = \frac{(\hat{p}-p_0)^2}{\hat{p}(1-\hat{p})/n}$, it has an asymptotic χ^2 distribution with df equal to the number of restrictions that are tested (Wooldridge 2012). This test is used for the purposes of testing Hypothesis 1.a to Hypothesis 6.e.

The other most frequently used possibility is the *the LR test* (likelihood-ratio test) that was already performed to exclude redundant variables from the model in Section 5.1. The LR test is based on the difference in the log-likelihood functions for the unrestricted and restricted models, where dropping a variable leads to a smaller or the same log-likelihood, due to the fact that $\mathcal{L}_{ur} \geq \mathcal{L}_r$, where $\mathcal{L}$ is the log-likelihood (Wooldridge 2012). LR statistic can be computed as $LR = 2(\mathcal{L}_{ur} - \mathcal{L}_r)$, having asymptotically χ^2_q, where q represent the number of df that is the number of restrictions in the restricted model.

5.2.1   Summary of hypotheses results

The Table 5.2 summarizes the results.

Table 5.2: Summary of hypotheses

Hypothesis	Sign. level	Result
1.a Device impacts WTP*	-	-
1.b OnlineNewsUse impacts WTP	0.345	Not supported
1.c PrintNewsUse impacts WTP	0.695	Not supported
2.a InternetUse impacts WTP	0.092	Partially supported
2.b NewsUse impacts WTP	0.006	Supported
3.a PastPayment impacts WTP	0.056	Partially supported
3.b PastNewspaperPayment impacts WTP	0.000	Supported
4.a Interest impacts WTP*	-	-
4.b Reader impacts WTP*	-	-
5.a LessAdd impacts WTP*	-	-
5.b CustomContent impacts WTP*	-	-
5.c QuealityContent impacts WTP*	-	-
6.a Male impacts positively WTP	0.006	Supported
6.b Age impacts negatively WTP	0.515	Not supported
6.c Educ impacts positively WTP	0.7613	Not supported
6.d Students more WTP	0.066	Partially supported
6.e Income impacts WTP	0.458	Not supported
7 Consumers more WTP for online	0.000	Supported

* hypothesis not tested in this section because variables were ruled out of
the model by LR tests, see Section B.2 in Appendix B

5.2.2 Hypotheses 1.a to 6.e

Coefficients *Device*, *LessAdd*, *CustomContent*, *QualityContent*, *Access*, and
Content were tested by the likelihood-ratio tests in Section 5.1. The null hy-
pothesis that the coefficient for these variables is equal to 0 cannot be rejected
even at 10 percent level of significance, therefore Hypothesis 1.a, Hypothe-
sis 5.a, Hypothesis 5.b, and Hypothesis 5.c cannot be supported even at 10
percent level of significance. Moreover, Hypothesis 4.a is not supported due to
the fact that the newly created variable *Content* (categorical variable taking
values of "none", "news", "tabloid") showed no joint significance when added
to estimated Equation 5.1. Therefore Hypothesis 4.b also cannot be supported,
as the hypothesis is based on the same variable *Content*. For likelihood-ratio
tests please see Table B.19 in Appendix B.

The remaining hypotheses are tested using the Wald test on the model (5.1)
estimated using the logit.

Hypotheses 1.b to 1.c

Hypothesis 1.b suggests that consuming more online news content positively influences the probability of having paying intentions. Positive relationship is already established, as the coefficient for $OnlineNewsUse$ is positive (see model (3) in Table 5.1), the marginal effect is therefore also positive as was established in Section 3.2 since $\frac{\partial P(\boldsymbol{x})}{\partial x_j} = g(\beta_0 + \boldsymbol{\beta x})\beta_j$ is purely dependent on the sign of β_j. Therefore we further need to know whether the coefficient is significant, thus setting $H_0 : coef(OnlineNewsUse) = 0$ against the alternative $H_a : coef(OnlineNewsUse) \neq 0$.

Table 5.3: Hypothesis 1.b, 1.c

(Intent)OnlineNewsUse = 0	chi2(1) = 0.89
	Prob > chi2 = 0.3453
(Intent)PrintNewsUse = 0	chi2(1) = 0.15
	Prob > chi2 = 0.6948

From Table 5.3 it is apparent that we cannot reject the null, since p-value is about 0.3453. Therefore the estimated coefficient for $OnlineNewsUse$ is not statistically significant even at 10 percent level of significance and Hypothesis 1.b cannot be supported. Similarly, Hypothesis 1.c implies that the consumption of printed news content have an effect on the paying intentions. Yet the coefficient for $PrintNewsUse$ is not significant and therefore Hypothesis 1.c cannot be supported even at 10 percent level of significance.

Hypotheses 2.a to 2.b

Hypothesis 2.a implies the relationship between the paying intentions and the time spend online. Hypothesis 2.b then suggests the relationship between paying intentions and the time spend on reading news. Again, single Wald tests were performed with following results.

Table 5.4: Hypothesis 2.a, 2.b

(Intent)InternetUse = 0	chi2(1) = 2.83
	Prob > chi2 = 0.0924
(Intent)NewsUse = 0 = 0	chi2(1) = 7.65
	Prob > chi2 = 0.0057

From Table 5.4 it is apparent that the coefficient for *InternetUse* is not equal to zero at 10 percent level of significance, therefore showing some support to the Hypothesis 1.a. On the other hand, variable *NewsUse* is significant at 1 percent level of significance, showing strong support to the Hypothesis 2.b. Therefore consumers of news content are more likely to be willing to pay for online news content (marginal effect of *NewsUse* is positive).

Hypotheses 3.a to 3.b

Hypothesis 3.a suggests that previous payments online (the variable *PastPayment*) in the past 3 months can influence the consumers' paying intentions. Similarly, Hypothesis 3.b suggests that those who have bought newspapers in the past 3 months show paying intentions for the print and therefore might share these intentions for the online content too.

Table 5.5: Hypothesis 3.a, 3.b

(Intent)PastPayment = 0	chi2(1) = 3.64
	Prob > chi2 = 0.0563
(Intent)0b.PastNewspaperPayment = 0	
(Intent)1.PastNewspaperPayment = 0	
(Intent)2.PastNewspaperPayment = 0	chi2(3) = 30.58
(Intent)3.PastNewspaperPayment = 0	
	Prob > chi2 = 0.0000

From Table 5.5 it is apparent that coefficient for the variable *PastPayment* is statistically different from 0 at 10 percent level of significance and also the marginal effect is positive, providing us with some support for Hypothesis 3.a. Furthermore, coefficients for the categorical variable *PastNewspaperPayment* are jointly significant at 1 percent level of significance, thus supporting Hypothesis 3.b that previous purchases of newspapers have a significant impact on consumers' WTP. The impact is also positive, as all three categories (1.Income for 1-3 times, 2.Income for 4-10 times, 3.Income for 10+ times) are marginally more probable to be WTP for the online content than those who have not bought a single copy of newspapers in the past 3 months (0.Income for 0 times). To see all the marginal effects please see Table B.22 in Appendix A.

Hypotheses 6.a to 6.e

Hypothesis 6.a suggests that men a have higher WTP than women. Therefore the coefficient for the variable $Male$ is expected to be statistically significant and positive. The coefficient for males is in fact positive, equal to 1.283, and statistically significant at 1 percent level of significance (see Table 5.6). Therefore the Hypothesis 6.a that men are more likely to pay for online content is supported. On the contrary, Hypothesis 6.b and Hypothesis 6.c cannot be supported due to the fact that respective coefficients are not statistically significant even at 10 percent level of significance

Table 5.6: Hypothesis 6.a, 6.b, 6.c

(Intent)Male = 0	chi2(1) = 7.44
	Prob > chi2 = 0.0064
(Intent)Age = 0	chi2(1) = 0.42
	Prob > chi2 = 0.5146
(Intent)Educ = 0	chi2(1) = 0.09
	Prob > chi2 = 0.7613

Table 5.7: Hypothesis 6.d, 6.e

(Intent)1b.Occup = 0	
(Intent)2.Occup = 0	
(Intent)3.Occup = 0	chi2(3) = 7.19
(Intent)4.Occup = 0	
	Prob > chi2 = 0.0660
(Intent)1b.Income = 0	
(Intent)2.Income = 0	
(Intent)3.Income = 0	chi2(3) = 2.60
(Intent)4.Income = 0	
	Prob > chi2 = 0.4582

Hypothesis 6.d suggests that students are more likely to pay for online content than employed readers. The categorical variable $Occupation$ needs to be jointly significant and marginal probability of students paying for online content has to be higher than of other groups. As illustrated in Table 5.7, the variable $Occupation$ is jointly significant at 10 percent level of significance. Since the base (1b.Occup) represents the students category, it needs to be marginally more probable to have paying intentions than employed category

(2.Occup). This is the case since as illustrated in Table B.22, the probability of an employed reader to have paying intentions is almost 6 percent lower than for students. Therefore there is some evidence to support Hypothesis 6.d.

Additionally, Hypothesis 6.e that consumers' WTP depends on their Income is represented by the jointly insignificant categorical variable *Income* even at 10 percent level of significance (see Table 5.7), providing no support for Hypothesis 6.e.

5.2.3   Hypothesis 7

Hypothesis 7 suggests that consumers would pay for the digital access to the news content less than for the printed counterpart. Table 5.8 summarizes the frequencies along with means and standard deviations.

Table 5.8: How much are people WTP for their favourite newpaper

Variable	Print	Online (Web&Application)
CZK	%	%
0-100	74.5	84.8
101-200	10.6	10.6
201-300	8.0	3.8
301+	6.8	6.8
Mean	96.7	50.7
SD	149.3	104.0
$N = 263$		

To be able to support this hypothesis, the mean value for the online access would arguably be lower than the mean value for the access to the printed edition.

To test this hypothesis *a two-sample t-test* with unequal variances is performed in Stata. The null hypothesis is that there is no difference in means between the two samples, such that $H_0 : \mu_{OnlineSubsription} - \mu_{PrintSubsription} = 0$, and an alternative that the difference is less than 0. The one-tailed test results in p-value equal to 0.000, suggesting a strong rejection of the null against this sided alternative.

Therefore it is apparent that consumers would pay significantly more for the printed edition when compared to the online edition. We conclude that Hypothesis 7 is supported.

Table 5.9: Hypothesis 7

$H_0 : \mu_{OnlineSubsription} - \mu_{PrintSubsription} = 0$	Ha: diff < 0
	$\Pr(\text{T} < \text{t}) = 0.0000$

5.3 Discussion

This study developed and validated a model to measure consumers' WTP for the online content. In particular, the study confirmed the perceived inferiority of the online content when compared to an alternative, in this case in form of printed newspapers, with p-value equal to 0. It has also showed that the WTP can be predicted using the logit model, that is fitted well with the outlined predictors.

Apparently, men in the Czech Republic are more willing to pay for online newspapers and subsequently for online content in general. This can be due to the fact that men use more interactive features of the websites and therefore can fully appreciate what is offered to them by online publishers. A suggestion might come out of this supported hypothesis that revenue from online sales can be increased by focusing more on content for men. Therefore this hypothesis has been confirmed in the Czech Republic by this thesis and in other regions by different researchers (Chung 2008; Chyi 2012).

People with a different occupation have different reading habits and different paying intentions. The students category may have been overestimated due to the collected sample properties, yet it seems that readers from this category show highest paying intentions for the online content. This can be correlated with the fact that younger people are more likely to be engaged in using newest technologies. Students are also well educated to realize what monetary value should they assign to a service of a given quality. On the other hand, the results were not that conclusive with the p-vlaue of 0.066, still showing some support to the hypothesized statement that students have higher paying intentions.

Internet usage of consumers also shows some significance. Intuitively this ought to be directly correlated with paying intentions if consumers are 'exposed' to the product—the online content, a lot, they will get used to it and consequently may pay for it. Alternatively, some users might use the Internet to just get the content and read it during, for example, travelling to work, and therefore these users would not spend a great amount of time using the Internet itself, yet might exhibit paying intentions.

Whether consumers did pay online in the past has an impact on the WTP. Further testing needs to be provided, because the support is not that conclusive with p-value of 0.056. Yet intuitively, this variable plays a role of a huge importance, because it reflects both the convenience and the lack of security concerns of consumers. It may be better to explore it in more details than with just a binary-response question that was used for the purposes of this thesis. Nevertheless, still having an impact on the estimated likelihood of having paying intentions.

Payments online are not the only type of payments that do affect the WTP. Whether consumers exhibit paying intentions online is significantly correlated with the paying behaviour for the printed news content. This means that if someone is used to buy newspapers regularly, he or she will not see a problem with spending their money for the same content, yet in a different form. On one hand, this may imply that the target audience for online newspapers would be readers of the print, yet other factors such as the perceived inferiority of the online content and the declining advertising revenue from selling less printed newspapers do occur. Therefore even though this factor has to be kept in mind when setting up a payment strategy for online publishers, for traditional printed newspapers the factor has to be evaluated cautiously.

The last of the significant factors that have an impact on paying intentions is the time spend reading the news content in any form. This implication is rather intuitive and is well supported at 1 percent level of significance with p-value of 0.006.

Apparently, the ownership of an electronic reading device with the Internet access does not affect the paying intentions. This is despite the fact that the Internet usage on its own does have an impact on the WTP. It is argued that one of the main advantages of online content is the convenience that is even higher with the usage of a transportable device. Yet this implication was not supported in the collected sample.

Whether consumers have read the online newspapers or digital newspapers in the previous three days have also showed no significance in the collected sample. This can imply that the collected variables did not carry a sufficient information that would significantly impact the WTP. Further researchers should take this into account and consider altering the question that is asked, in order to identify the desired impact of the observed variable.

Interest in a newspaper content and the actual reading habit of a newspaper content were correlated in the studied sample. Even after ruling out the read-

ing habit variable and simplifying the Interest variable into a single variable *Content*, no significance was shown in the estimated model. On one hand, it can be a property of the collected sample. On the other, a better design of the studied variable is advised for future researchers.

It proved to be difficult to properly reflect the expectations for paid online content in the WTP. All three of the studied variables did not improve the estimated likelihood of having paying intentions. Still, statistics show that the most valued property of online content is its quality (mean value 4.47, while 5 corresponds to the 'Strongly agree' option), the second one is the amount of advertising that seems to be not desirable (mean value 4.11), and the least valued is the ability to customize the content to consumers' needs (mean value 3.72). For frequencies please see Appendix B.

Other factors showed no impact on WTP for online content, such as age, education, or income. These variables are suggested to be studied further, because of the contradictory results of previous studies (Chyi 2005; BCG 2009; OPA 2013; Goyanes 2014; WAN 2014).

5.4 Limitations

Studies analysing WTP face limitations for several reasons, as stated in Section 2.2. Regarding this thesis, the first limitation to be pointed out is the possibility that some of the factors influencing the WTP for online content in the Czech Republic were not properly reflected in the analysis or even identified due to the lack of a previous occurrence in the relevant literature (such as the importance of a brand, more detailed security concerns, or reader types). This may be because the human behaviour is difficult to predict and broken down into factors that influence it the most, and that the area of WTP is not yet fully understood. Yet this thesis tried to use both the previously used factors and those arising from intuition.

It seems that some of the factors were not identified properly or were not sufficiently reflected in the collected sample. The collected sample may also not have been randomly selected from the whole Internet population, because the majority of respondents were from the Author's circle of acquaintances. It can also suffer from a self-selection bias due to the fact that only those who wanted to complete the questionnaire did so. The sample consisted mostly from students, having the average age rather low and educational level high. More questions arose during writing this thesis that have not been considered when

outlining the research. The questionnaire included two questions which answers were directly correlated even though the intuition might have suggested otherwise. Also other methodology is to be considered, such as using the ordered logit to estimate the effects for the full likert scale, yet the responses would have to be distributed in a way to allow for it.

The most importantly, we have to keep in mind that the research was designed in such a way that only the hypothesized preferences were collected. These preferences may not translate into an actual behaviour when put in front of making a real decision. Nevertheless, the studied topic would be difficult to observe in an experiment environment, therefore researchers have to rely on these stated preferences or a real data reflecting the actual behaviour of consumers.

Chapter 6

Conclusion

This thesis studied the research question: **Are inhabitants in the Czech Republic willing to pay for online content?** Relying on statistics, only about 2 percent of individuals in the Czech Republic used the Internet to order or buy the online content in 2014, which is bellow the European average of 15 percent (Eurostat 2015). Most inhabitants in the Czech Republic would therefore not pay for the online content, neither for digital news content in particular.

A more detailed answer consists of identifying factors that influence the probability of having paying intentions. Using the logit model, the following factors were identified to have a significant impact on the dependent variable paying intent. First of these variables is *NewsUse* that measured the amount of time consumers spend on reading newspapers. The next variable *InternetUse* measured the amount of time consumers spend on using the Internet and it has also showed a significant relationship with paying intentions. Other than that, the gender variable showed that men are more WTP for the online news content than women. What also influences the paying intentions is the previous paying behaviour. The variable *PastPayment* indicated whether consumers did pay using the Internet in previous 3 months and has showed significance. *PastNewspaperPayments* also positively impacts the WTP for the online content, representing the amount of times consumers bought a printed copy of newspapers in the past 3 months. These results are consistent with the relevant literature. According to Chyi (2012), men are more WTP for online content, along with the direct relation of the time spend on reading news and the paying intentions. Moreover, Internet use also affects WTP as argued by Lin *et al.* (2013). Goyanes (2014) and Dou (2004) previously suggested the importance

of the past online payments and also the direct payments for the content of interest, well applicable to the newspaper industry.

One of the reasons for the unwillingness for online content is that Internet users assign the digital news content with a lower monetary value than they do in the case of printed newspapers. This argument was broadened to a hypothesized scenario when there are no free substitutes, that would provide consumers with online news content at no charge. Nevertheless, on average Internet users would pay CZK 96.7 for a printed subscription of their favourite newspaper and only CZK 50.7 for the online (access to the application and the website) alternative. The difference showed statistical significance, therefore it is safe to conclude, that inhabitants in the Czech Republic in early 2015 still find the online news content to be inferior, or at least less valuable, when compared to the printed alternative. This further confirms the perceived inferiority of online content suggested by Chyi & Yang (2009).

Although the process of finding the audience willing to pay for online content is rather slow, online publishers have managed to attract some of Internet readers even in the Czech Republic. Globally, digital advertising and sales revenues are steadily increasing (WAN 2014), proving that the Internet is a suitable environment to generate revenue. This along with the increasing usage of electronic reading devices (PEW 2013) helps to create a new, dynamic environment with a potential that is yet to be unlocked.

Bibliography

ADAMS, R. (2012): "Papers Put Faith in Paywalls: Digital Subscription Sales Are Slow, but Publishers Are Encouraged by Trends." *The Wall Street Journal* `http://www.wsj.com/articles/SB10001424052970203833004577251822631536422`, Accessed: 2015.15.03.

ADDISON, D. (2001): "Free web access business model is unsustainable in the long term." *Marketing* pp. 9–10.

BAKKER, P. (2002): "Free daily newspapers-business models and strategies." *International Journal on Media Management* **4(3)**: pp. 180–187.

BAKKER, P. (2007): "Free newspaper readership." In "Worldwide Readership Research Symposia Vienna 2007," pp. 63–74. WRRS.

BAKKER, P. (2008): "The simultaneous rise and fall of free and paid newspapers in europe." *Journalism Practice* **2(3)**: pp. 427–443.

BCG (2009): "Who would pay for online news: Key findings from an international survey." *The Boston Consulting Group* `https://www.alternatives-economiques.fr/fic_bdd/article_pdf_fichier/1261153757_BCG_online_news_survey-France-20Nov09.pdf`, Accessed: 2015.15.03.

BLEYEN, V.-A. & L. VAN HOVE (2007): "Western european newspapers and their online revenue models: An overview." *First Monday (The Peer-Reviewed Journal on the Internet)* **12(12)**.

BREIDERT, C., M. HAHSLER, & T. REUTTERER (2006): "A review of methods for measuring willingness-to-pay." *Innovative Marketing* **2(4)**: pp. 8–32.

CARLSON, D. (2003): "The history of online journalism." *Digital journalism: Emerging media and the changing horizons of journalism* pp. 31–55.

CHUNG, D. S. (2008): "Interactive features of online newspapers: Identifying patterns and predicting use of engaged readers." *Journal of Computer-Mediated Communication* **13(3)**: pp. 658–679.

CHYI, H. I. (2005): "Willingness to pay for online news: An empirical study on the viability of the subscription model." *Journal of Media Economics* **18(2)**: pp. 131–142.

CHYI, H. I. (2012): "Paying for what? how much? and why (not)? predictors of paying intent for multiplatform newspapers." *International Journal on Media Management* **14(3)**: pp. 227–250.

CHYI, H. I. & M. CHADHA (2012): "News on new devices: Is multi-platform news consumption a reality?" *Journalism Practice* **6(4)**: pp. 431–449.

CHYI, H. I. & D. L. LASORSA (2002): "An explorative study on the market relation between online and print newspapers." *The Journal of Media Economics* **15(2)**: pp. 91–106.

CHYI, H. I. & A. M. LEE (2012): "Theorizing online news consumption: A structural model linking preference, use, and paying intent." In "13th International Symposium on Online Journalism, Austin, TX, April," pp. 20–21.

CHYI, H. I. & M. J. YANG (2009): "Is online news an inferior good? examining the economic nature of online news among users." *Journalism & Mass Communication Quarterly* **86(3)**: pp. 594–612.

DCN (2014): "Pew offers insights into smartphone users and usage." *Research Team DCN* http://digitalcontentnext.org/blog/2015/04/06/pew-offers-insights-into-smartphone-users-and-usage/, Accessed: 2015.15.03.

DEWAN, R. M., M. L. FREIMER, & J. ZHANG (2002): "Management and valuation of advertisement-supported web sites." *Journal of Management Information Systems* **19(3)**: pp. 87–98.

DIMMICK, J., Y. CHEN, & Z. LI (2004): "Competition between the internet and traditional news media: The gratification-opportunities niche dimension." *The Journal of Media Economics* **17(1)**: pp. 19–33.

DOU, W. (2004): "Will internet users pay for online content?" *Journal of Advertising Research* **44(04)**: pp. 349–359.

DUTTA-BERGMAN, M. J. (2004): "Complementarity in consumption of news types across traditional and new media." *Journal of Broadcasting & Electronic Media* **48(1)**: pp. 41–60.

EDMONDS, R. (2009): "Is Paid Online Content a Solution or Impossible Dream?" *Poynter* `http://www.poynter.org/news/mediawire/94173/is-paid-online-content-a-solution-or-impossible-dream/`, Accessed: 2015.15.03.

ELVESTAD, E. & A. BLEKESAUNE (2008): "Newspaper readers in europe a multilevel study of individual and national differences." *European Journal of Communication* **23(4)**: pp. 425–447.

EUROSTAT (2015): "Individuals using the internet to buy or order online content." *Eurostat* `http://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&plugin=1&pcode=tin00080&language=en`, Acessed: 2015.15.03.

FARHI, P. (2008): "Don't blame the journalism." *American Journalism Review* **30(5)**: pp. 14–15. `http://ajrarchive.org/Article.asp?id=4623`, Accessed: 2015.15.03.

FISHER, L. (2005): "Random sampling versus panel-based research." *MRA Alert Magazine* .

GOYANES, M. (2014): "An empirical study of factors that influence the willingness to pay for online news." *Journalism Practice* **8(6)**: pp. 742–757.

HANEMANN, W. M. (1991): "Willingness to pay and willingness to accept: how much can they differ?" *The American Economic Review* pp. 635–647.

HERBERT, J. & N. THURMAN (2007): "Paid content strategies for news websites: An empirical study of british newspapers' online business models." *Journalism Practice* **1(2)**: pp. 208–226.

HOFFMAN, E., D. J. MENKHAUS, D. CHAKRAVARTI, R. A. FIELD, & G. D. WHIPPLE (1993): "Using laboratory experimental auctions in marketing research: A case study of new packaging for fresh beef." *Marketing science* **12(3)**: pp. 318–338.

ITU (2014): "Facts and Figures." `http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2014-e.pdf`, Accessed: 2015.15.03.

KAGEL, J. & D. LEVIN (2002): "Bidding in common-value auctions: A survey of experimental research." *Common value auctions and the winner's curse* **1**: pp. 1–84.

KOHUT, A., C. DOHERTY, M. DIMOCK, & S. KEETER (2010): "Americans spending more time following the news." *Pew Research Center* .

KOPITS, I. M., C. CHEN, J. S. ROBERTS, W. UHLMANN, & R. C. GREEN (2011): "Willingness to pay for genetic testing for alzheimer's disease: a measure of personal utility." *Genetic testing and molecular biomarkers* **15(12)**: pp. 871–875.

LAUF, E. (2001): "Research note: The vanishing young reader sociodemographic determinants of newspaper use as a source of political information in europe, 1980-98." *European Journal of Communication* **16(2)**: pp. 233–243.

LIN, T.-C., J. S.-C. HSU, & H.-C. CHEN (2013): "Customer willingness to pay for online music: The role of free mentality." *Journal of Electronic Commerce Research* **14(4)**: pp. 315–333.

LOPES, A. B. & D. F. GALLETTA (2006): "Consumer perceptions and willingness to pay for intrinsically motivated online content." *Journal of Management Information Systems* **23(2)**: pp. 203–231.

MARN, M. V., E. V. ROEGNER, & C. C. ZAWADA (2003): "Pricing new products." *McKinsey Quarterly* **(3)**: pp. 40–49.

MEDIAPROJEKT (2013): "Online Nadstavba Media Projektu." `http://www.unievydavatelu.cz/gallery/files/NastavbaMP_elektronicke_platformy_tisku_leden%202014.pdf`, Accessed: 2015.15.03.

MINDICH, D. T. (2005): *Tuned out: Why Americans under 40 don't follow the news.* Oxford University Press.

MINGS, S. M. & P. B. WHITE (2000): "Profiting from online news: The search for viable business models." *Internet publishing and beyond: The economics of digital information and intellectual property* pp. 62–96.

OPA (2013): "Digital content subscription strategies pay off for publishers." *Pew Research Center* .

PEW (2013): "In changing news landscape, even television is vulnerable." *PewResearchCenter* `http://www.people-press.org/2012/09/27/in-changing-news-landscape-even-television-is-vulnerable/`, Accessed: 2015.15.03.

PICARD, R. G. (2001): "Strategic responses to free distribution daily newspapers." *International Journal on Media Management* **3(3)**: pp. 167–172.

POINDEXTER, P., S. MERAZ, & A. S. WEISS (2010): *Women, men and news: Divided and disconnected in the news media landscape.* Routledge.

PORTER, M. E. (2001): "Strategy and the internet." *Harvard Business Review* **79(3)**: pp. 62–78.

PRIOR, M. (2007): *Post-broadcast democracy: How media choice increases inequality in political involvement and polarizes elections.* Cambridge University Press.

REUTERS (2014): "Digital news report." *Reuters Institute for the Study of Journalism* .

SCHIFFMAN, B. (2001): "Free Web Services: There's No Turning Back." *Forbes* `http://www.forbes.com/2001/03/20/0320free.html`, Accessed: 2015.15.03.

SCHONFELD, E. (2001): "Future boy: The napster legacy." *Business* **2**: pp. 51–53.

SCHWER, R. K. & R. DANESHVARY (1995): "Willingness to pay for public television and the advent of" look-alike" cable television channels: A case study." *Journal of Media Economics* **8(3)**: pp. 95–109.

SHOGREN, J. F., S. Y. SHIN, D. J. HAYES, & J. B. KLIEBENSTEIN (1994): "Resolving differences in willingness to pay and willingness to accept." *The American Economic Review* pp. 255–270.

SILLANO, M. & J. DE DIOS ORTÚZAR (2005): "Willingness-to-pay estimation with mixed logit models: some new evidence." *Environment and Planning* **37**: pp. 525—550.

SMITH, S. (2010): "Are We Ready to Play With Pay?" `http://www.econtentmag.com/Articles/Editorial/Feature/Are-We-Ready-to-Play-With-Pay-The-Content-Value-Reproposition-66000.htm`, Accessed: 2015.15.03.

TOM ROSENSTIEL, Mark Jurkowitz, H. J. (2012): "The Search for a New Business Model: How Newspapers are Faring Trying to Build Digital Revenue." *PewResearchCenter* `http://www.journalism.org/2012/03/05/search-new-business-model/`, Accessed: 2015.15.03.

VICKREY, W. (1961): "Counterspeculation, auctions, and competitive sealed tenders." *The Journal of finance* **16(1)**: pp. 8–37.

WAN (2014): "World Press Trends." *World Association of Newspapers and News Publishers* `http://www.arpp.ru/images/123/51253_WAN-IFRA_WPT_2014.pdf`, Accessed: 2015.15.03.

WANG, C. L., L. R. YE, Y. ZHANG, & D.-D. NGUYEN (2005): "Subscription to fee-based online services: what makes consumer pay for online content." *Journal of Electronic Commerce Research* **6(4)**: pp. 304–311.

WERTENBROCH, K. & B. SKIERA (2002): "Measuring consumers' willingness to pay at the point of purchase." *Journal of Marketing Research* **39(2)**: pp. 228–241.

WOOLDRIDGE, J. (2012): *Introductory econometrics: A modern approach.* Cengage Learning.

YOUNG, M. R., W. S. DESARBO, & V. G. MORWITZ (1998): "The stochastic modeling of purchase intentions and behavior." *Management Science* **44(2)**: pp. 188–202.

ZEITHAML, V. A. (1988): "Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence." *The Journal of marketing* pp. 2–22.

Appendix A

Online questionnaire

A.1   Online questionnaire

1. Kolik hodin denně strávíte na Internetu? (Př. hodina a půl, vložte "1.5") *
How many hours do you spend using the Internet per day? (Eg. an hour and half, insert "1.5")

2. Vlastním smartphone/tablet/čtečku s přístupem k Internetu. *
I have a smartphone/tablet/e-reader with the Internet access.

- ◯ Ano (Yes)
- ◯ Ne (No)

3. Jak se nejčastěji připojujete k Internetu *
How do you mostly access the Internet?

- ◯ Z práce/ze školy (Work/school)
- ◯ Z osobního počítače (Personal computer)
- ◯ Ze smartphonu/tabletu/čtečky (Smartphone/tablet/e-readder)
- ◯ Jiné (Other)

4. Platil/a jste v posledních 3 měsících prostřednictvím Internetu? *
Have you paid using the Internet in past 3 months?

- ◯ Ano (Yes)
- ◯ Ne (No)

5. Během posledních 3 dní jsem četl... *
During the past 3 days I have read...

- ☐ Tištěné noviny (Hospodářské noviny, MF Dnes, Lidové noviny, Právo, Aha!, Blesk, Sport, ...) (Print)
- ☐ Online noviny (novinky.cz, idnes.cz, ihned.cz, aktualne.cz, blesk.cz, super.cz, sport.cz, ...) (Online)
- ☐ Žádné (None)

6. Kolik hodin denně strávíte čtením novin (online+tisk)? (Př. hodina a půl, vložte "1.5") *
How many hours do you spend reading newspapers per day (online+print)? (Eg. an hour and half, insert "1.5")

☐

7. Koupil jste si za poslední tři měsíce noviny? *
Have you purchased newspapers in the past 3 months?

◯ Ne (No)
◯ 1 - 3 krát (1 - 3 times)
◯ 3 - 10 krát (3 - 10 times)
◯ Vícekrát než 10 krát (More than 10 times)

8. Nejčastěji čtu... *
Most often I read...

◯ Zpravodajství (Hospodářské noviny, MF Dnes, Lidové noviny, Právo, idnes.cz, ihned.cz, aktualne.cz ...) (News)
◯ Společenské dění (Aha!, Blesk, blesk.cz, super.cz, ...) (Tabloid)
◯ Sport (Sport, sport.cz, ...)
◯ Specializovaný obsah (ekonomika, pro ženy, zábava a hry, technologie, auto-moto, ...) (Specialized content)
◯ Žádné (None)

9. Nejvíce se zajímám o... *
I am mostly interested in...

◯ Zpravodajství (Hospodářské noviny, MF Dnes, Lidové noviny, Právo, idnes.cz, ihned.cz, aktualne.cz ...) (News)
◯ Společenské dění (Aha!, Blesk, blesk.cz, super.cz, ...) (Tabloid)
◯ Sport (Sport, sport.cz, ...)
◯ Specializovaný obsah (ekonomika, pro ženy, zábava a hry, technologie, auto-moto, ...) (Specialized content)
◯ Žádné (None)

10. Některé noviny zvažují zpoplatnění svého online obsahu. Jak moc byste osobně byl/a ochoten/a platit za takový obsah? *
Some newspapers are considering to charge users for online content. How likely is it that you personally would pay for news and information on their website?

	1	2	3	4	
Určitě ne (Very unlikely)	◯	◯	◯	◯	Určitě ano (Very likely)

Od placených online novin očekávám...

I expect from the paid online news content

11. Méně/žádná reklama *
Less/No advertising

1 2 3 4 5

Vůbec nesouhlasím (Strongly disagree) ◯ ◯ ◯ ◯ ◯ Hodně souhlasím (Strongly agree)

12. Obsah přizpůsobený uživateli *
Customized content

1 2 3 4 5

Vůbec nesouhlasím (Strongly disagree) ◯ ◯ ◯ ◯ ◯ Hodně souhlasím (Strongly agree)

13. Kvalitní a podrobný obsah *
Verified detailed content

1 2 3 4 5

Vůbec nesouhlasím (Strongly disagree) ◯ ◯ ◯ ◯ ◯ Hodně souhlasím (Strongly agree)

Představte si, že Vaše oblíbené noviny již nejsou nikde zadarmo (ani tištěné, ani online). V takovém případě, kolik byste byli ochotni platit souběžně za Vaše oblíbené noviny na jednotlivých platformách? Můžete si vybrat jednu variantu, obě, nebo žádnou. Pokud nic, vložte 0. [Kč]

Imagine that no newspaper content is free anymore. How much would you pay for respective reading platforms of your favourite newspapers? You can choose one option, both, or none. If not anything, put 0. [Czech Korunas]

14. Tištěná edice (měsíční předplatné) *
Printed edition (monthly subscription)

[   ]

15. Online edice (měsíční přístup na web, do aplikace) *
Online edition (monthly access to web platform, application)

[   ]

Demografické otázky
Demographic questions

16. Jaké je Vaše pohlaví? *
What is your gender?

◯ Muž (Male)

◯ Žena (Female)

17. Kolik je Vám let? *
How old are you?

18. Jaké jste národnosti? *
What is your nationality?
- Česká (Czech)
- Slovenská (Slovak)
- Jiná (Other)

19. Jakého nejvyššího vzděláni jste dosáhl/la? *
What is your achieved level of education?
- Základní (Elementary)
- Středoškolské (High school)
- Vysokoškolské (University)

20. Jsem... *
What is your occupation?
- Student
- Zaměstnaný (Employed)
- OSVČ (Self-employed)
- Nezaměstnaný (Unemployed)

21. Jaký je Váš čistý měsíční příjem? [Kč] *
What is your monthly after-tax income? [Czech Korunas]
- 0 - 10 000
- 10 001 - 20 000
- 20 001 - 30 000
- 30 001 - 50 000
- 50 001+

Submit

Appendix B

Results

This section of the appendix consists of various results from different models, but also from the frequency tables. Since a lot of the variables are categorical, a series of dummies were used for the purposes of the estimation. Dummies were created by the *Stata 12.0* software in such a way to correspond to the outlined hypotheses. For more details on the coding of the variables, please see Section 4.2, categorical variables are listed below. The *Stata* software is also responsible for all the computations. The most of the results are copied directly from *Stata* using the `estout` package.

Categorical variables codes

- *Occupation*, a categorical variable with 1 for "student", 2 for "employed", 3 for "self-employed", and 4 for "unemployed".

- *Income* a categorical variable with 1 for "0-10 thousands Czech Korunas after tax per month", 2 for "10-20", 3 for "20-30", 4 for "30+".

- *Access*, categorical variable with 1 for "from a personal computer", 2 for "from a mobile device", 3 "from work/school" and 4 "from elsewhere".

- *PastPaymentNewspapers*, a categorical variable with 0 for "bought newspapers 0 times in the past 3 months", 1 for "1-3 times", 2 for "4-6 times" and 3 for "10+ times".

- *Content*, a categorical variable with 0 for "other content", 1 for "news content" and 2 for "specialized content".

B.1   Data description and frequency tables

Table B.1:   Age

Item	Number	Per cent
15	4	2
16	1	0
17	4	2
18	9	3
19	17	6
20	9	3
21	32	12
22	63	24
23	23	9
24	6	2
25	18	7
26	8	3
27	8	3
28	3	1
29	2	1
30	1	0
32	1	0
33	3	1
36	2	1
38	6	2
39	1	0
41	2	1
42	5	2
44	7	3
45	2	1
46	1	0
47	8	3
48	2	1
49	2	1
50	3	1
51	1	0
53	3	1
55	3	1
57	1	0
61	2	1
Total	263	100

Table B.2:   Male

Item	Number	Per cent
0	123	47
1	140	53
Total	263	100

Table B.3:   Income

Item	Number	Per cent
1	150	57
2	48	18
3	42	16
4	23	9
Total	263	100

Table B.4:   Educ

Item	Number	Per cent
9	16	6
13	162	62
16	85	32
Total	263	100

Table B.5:   Occup

Item	Number	Per cent
1	173	66
2	83	32
3	5	2
4	2	1
Total	263	100

Table B.6:   InternetUse

Item	Number	Per cent
0	2	1
.3	1	0
.5	8	3
1	23	9
1.4926	2	1
1.5	4	2
2	31	12
2.5	2	1
3	36	14
3.5	5	2
4	37	14
5	35	13
6	40	15
7	7	3
8	14	5
9	4	2
10	8	3
11	1	0
12	3	1
Total	263	100

Table B.7:   NewsUse

Item	Number	Per cent
0	25	10
.1	2	1
.1376	2	1
.15	1	0
.2	6	2
.25	5	2
.3	4	2
.5	97	37
.7	3	1
.75	2	1
.8	1	0
1	77	29
1.5	3	1
2	34	13
3	1	0
Total	263	100

Table B.8:   Device

Item	Number	Per cent
0	50	19
1	213	81
Total	263	100

Table B.9:   OnlineNewsUse

Item	Number	Per cent
0	35	13
1	228	87
Total	263	100

Table B.10:   PrintNewsUse

Item	Number	Per cent
0	198	75
1	65	25
Total	263	100

Table B.11:   PastPayment

Item	Number	Per cent
0	46	17
1	217	83
Total	263	100

Table B.12:   PastNewspaperPayment

Item	Number	Per cent
0	175	67
1	53	20
2	17	6
3	18	7
Total	263	100

Table B.13: Reader

Item	Number	Per cent
0	12	5
1	157	60
2	8	3
3	25	10
4	61	23
Total	263	100

Table B.14: Interest

Item	Number	Per cent
0	7	3
1	137	52
2	5	2
3	27	10
4	87	33
Total	263	100

Table B.15: Content

Item	Number	Per cent
0	87	33
1	137	52
2	39	15
Total	263	100

Table B.16: LessAdd

Item	Number	Per cent
1	17	6
2	19	7
3	30	11
4	49	19
5	148	56
Total	263	100

Table B.17: CustomContent

Item	Number	Per cent
1	9	3
2	20	8
3	83	32
4	74	28
5	77	29
Total	263	100

Table B.18: QualityContent

Item	Number	Per cent
1	4	2
2	5	2
3	25	10
4	58	22
5	171	65
Total	263	100

B.2   Likelihood-ratio tests

```
logistic Intent Age Male Educ i.Occup i.Income Czech
        estimates store m1
logistic Intent Age Male Educ i.Occup i.Income Czech  Device i.Access
        estimates store m2
        lrtest m1 m2
logistic Intent Age Male Educ i.Occup i.Income Czech OnlineNewsUse PrintNewsUse
        estimates store m3
        lrtest m1 m3
logistic Intent Age Male Educ i.Occup i.Income Czech OnlineNewsUse PrintNewsUse NewsUse
InternetUse
        estimates store m4
        lrtest m3 m4
logistic Intent Age Male Educ i.Occup i.Income Czech OnlineNewsUse PrintNewsUse NewsUse
InternetUse PastPayment i.PastNewspaperPayment
        estimates store m5
        lrtest m4 m5
logistic Intent Age Male Educ i.Occup i.Income Czech  OnlineNewsUse PrintNewsUse NewsUse
InternetUse PastPayment i.PastNewspaperPayment i.Content
        estimates store m6
        lrtest m5 m6
logistic Intent Age Male Educ i.Occup i.Income Czech  OnlineNewsUse PrintNewsUse NewsUse
InternetUse PastPayment i.PastNewspaperPayment LessAdd CustomContent QualityContent
        estimates store m7
        lrtest m5 m7
logistic Intent Age Male Educ i.Occup i.Income Czech  OnlineNewsUse PrintNewsUse NewsUse
InternetUse PastPayment i.PastNewspaperPayment i.Content LessAdd CustomContent QualityContent
        estimates store m8
        lrtest m5 m8
```

Table B.19: Likelihood-ratio tests - all results

Likelihood-ratio test (Assumption: m1 nested in m2)	LR chi2(3) = 1.86 Prob > chi2 = 0.6016
Likelihood-ratio test (Assumption: m1 nested in m3)	LR chi2(2) = 9.76 Prob > chi2 = 0.0076
Likelihood-ratio test (Assumption: m3 nested in m4)	LR chi2(2) = 23.70 Prob > chi2 = 0.0000
Likelihood-ratio test (Assumption: m4 nested in m5)	LR chi2(4) = 42.42 Prob > chi2 = 0.0000
Likelihood-ratio test (Assumption: m5 nested in m6)	LR chi2(2) = 2.12 Prob > chi2 = 0.3460
Likelihood-ratio test (Assumption: m5 nested in m7)	LR chi2(3) = 1.41 Prob > chi2 = 0.7036
Likelihood-ratio test (Assumption: m5 nested in m8)	LR chi2(5) = 2.89 Prob > chi2 = 0.7162

B.3 LPM

```
eststo: xi:  reg Intent Age Male Educ i.Occup i.Income Czech, vce(robust)
eststo: xi:  reg Intent Age Male Educ i.Occup i.Income Czech  Device i.Access, vce(robust)
eststo: xi:  reg Intent Age Male Educ i.Occup i.Income Czech OnlineNewsUse PrintNewsUse
NewsUse InternetUse PastPayment i.PastNewspaperPayment, vce(robust)
eststo: xi:  reg Intent Age Male Educ i.Occup i.Income Czech  Device i.Access OnlineNewsUse
PrintNewsUse NewsUse InternetUse PastPayment i.PastNewspaperPayment i.Content LessAdd
CustomContent QualityContent, vce(robust)
```

Table B.20: LPM vce(robust)

	(1) Intent	(2) Intent	(3) Intent	(4) Intent
Age	-0.00489	-0.00435	-0.00318	-0.00445
Male	0.158**	0.148*	0.136**	0.121
Educ	0.00118	0.00163	-0.00196	-0.00241
_IOccup_2	-0.278*	-0.277*	-0.0978	-0.0588
_IOccup_3	-0.0374	-0.0400	0.0863	0.130
_IOccup_4	0.449	0.489	0.482	0.524
_IIncome_2	0.200*	0.200	0.0858	0.0632
_IIncome_3	0.302*	0.291*	0.127	0.0934
_IIncome_4	0.508***	0.492***	0.275*	0.230
Czech	-0.0736	-0.0730	-0.0402	-0.0227
Device		0.0677		0.0258
_IAccess_2		-0.0145		-0.0671
_IAccess_3		0.0484		0.0568
OnlineNewsUse			0.0355	0.0246
PrintNewsUse			-0.0410	-0.0359
NewsUse			0.148**	0.149**
InternetUse			0.0210	0.0180
PastPayment			0.121*	0.102
_IPastNewsp_1			0.320***	0.330***
_IPastNewsp_2			0.485***	0.490***
_IPastNewsp_3			0.339**	0.339**
_IContent_1				-0.0178
_IContent_2				-0.0979
LessAdd				0.000722
CustomContent				-0.0255
QualityContent				0.0363
_cons	0.315	0.246	-0.143	-0.131
N	263	263	263	263
R^2	0.122	0.125	0.363	0.379

$^{*}\ p < 0.05,\ ^{**}\ p < 0.01,\ ^{***}\ p < 0.001$

B.4   Model comparison

```
eststo: xi: logit Intent Age Male Educ i.Occup i.Income Czech OnlineNewsUse PrintNewsUse
NewsUse InternetUse PastPayment i.PastNewspaperPayment
eststo: xi: probit Intent Age Male Educ i.Occup i.Income Czech OnlineNewsUse PrintNewsUse
NewsUse InternetUse PastPayment i.PastNewspaperPayment
eststo: xi: reg Intent Age Male Educ i.Occup i.Income Czech OnlineNewsUse PrintNewsUse
NewsUse InternetUse PastPayment i.PastNewspaperPayment, vce(robust)
```

Table B.21: Model comparison (logit, probit, lpm)

	(1) Logit Intent		(2) Probit Intent		(3) LPM Intent	
Age	-0.0263	(0.515)	-0.0156	(0.500)	-0.00318	(0.398)
Male	1.283**	(0.006)	0.724**	(0.005)	0.136**	(0.009)
Educ	0.0386	(0.761)	0.0357	(0.626)	-0.00196	(0.886)
_IOccup_2	-0.536	(0.497)	-0.380	(0.372)	-0.0978	(0.369)
_IOccup_3	0.551	(0.684)	0.246	(0.746)	0.0863	(0.721)
_IOccup_4	4.055*	(0.033)	2.345*	(0.043)	0.482	(0.165)
_IIncome_2	0.220	(0.699)	0.194	(0.539)	0.0858	(0.350)
_IIncome_3	0.535	(0.520)	0.326	(0.497)	0.127	(0.235)
_IIncome_4	1.643	(0.118)	0.984	(0.102)	0.275*	(0.030)
Czech	-0.503	(0.418)	-0.275	(0.445)	-0.0402	(0.593)
OnlineNewsUse	0.656	(0.345)	0.397	(0.329)	0.0355	(0.556)
PrintNewsUse	-0.196	(0.695)	-0.0836	(0.767)	-0.0410	(0.548)
NewsUse	1.021**	(0.006)	0.600**	(0.004)	0.148**	(0.009)
InternetUse	0.154	(0.092)	0.0931	(0.073)	0.0210	(0.133)
PastPayment	1.290	(0.056)	0.694*	(0.048)	0.121*	(0.040)
_IPastNewsp_1	2.201***	(0.000)	1.247***	(0.000)	0.320***	(0.000)
_IPastNewsp_2	3.224***	(0.000)	1.853***	(0.000)	0.485***	(0.000)
_IPastNewsp_3	2.468**	(0.002)	1.392**	(0.002)	0.339**	(0.005)
_cons	-5.751**	(0.006)	-3.481**	(0.004)	-0.143	(0.522)
N	263		263		263	
R^2					0.363	
pseudo R^2	0.361		0.363			

p-values in parentheses

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

B.5 Logit fit

```
logit Intent Age Male Educ i.Occup i.Income Czech OnlineNewsUse PrintNewsUse NewsUse
InternetUse PastPayment i.PastNewspaperPayment

lstest
```

```
          Logistic model for Intent


                      -------- True --------
          Classified |         D          ~D |        Total
          -----------+-------------------------+-----------
              +       |        40          16 |           56
              -       |        29         178 |          207
          -----------+-------------------------+-----------
           Total      |        69         194 |          263


          Classified + if predicted Pr(D) >= .5
          True D defined as Intent != 0
          --------------------------------------------------
          Sensitivity                   Pr( +| D)   57.97%
          Specificity                   Pr( -|~D)   91.75%
          Positive predictive value     Pr( D| +)   71.43%
          Negative predictive value     Pr(~D| -)   85.99%
          --------------------------------------------------
          False + rate for true ~D      Pr( +|~D)    8.25%
          False - rate for true D       Pr( -| D)   42.03%
          False + rate for classified + Pr(~D| +)   28.57%
          False - rate for classified - Pr( D| -)   14.01%
          --------------------------------------------------
          Correctly classified                      82.89%
          --------------------------------------------------
```

B.6   Marginal effects

```
logit Intent Age Male Educ i.Occup i.Income Czech OnlineNewsUse PrintNewsUse NewsUse
InternetUse PastPayment i.PastNewspaperPayment

margins, dydx(*) post
```

Table B.22: Marginal effects of logit model

	(1) dy/dx	
Age	-0.00304	(0.513)
Male	0.148**	(0.004)
Educ	0.00446	(0.761)
1b.Occup	.	.
2.Occup	-0.0598	(0.486)
3.Occup	0.0701	(0.697)
4.Occup	0.557**	(0.003)
1b.Income	.	.
2.Income	0.0250	(0.704)
3.Income	0.0633	(0.534)
4.Income	0.217	(0.140)
Czech	-0.0581	(0.417)
OnlineNewsUse	0.0758	(0.343)
PrintNewsUse	-0.0226	(0.694)
NewsUse	0.118**	(0.004)
InternetUse	0.0178	(0.087)
PastPayment	0.149	(0.051)
0b.PastNewspaperPayment	.	.
1.PastNewspaperPayment	0.312***	(0.000)
2.PastNewspaperPayment	0.484***	(0.000)
3.PastNewspaperPayment	0.357**	(0.005)
N	263	

p-values in parentheses

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