PRICE DISPERSION IN THE ONLINE BOOK MARKET

by

Bohdan Korzh

A thesis submitted in partial fulfillment of the
requirements for the degree of

MA in Business and Financial Economics

Kyiv School of Economics

2021

Thesis Supervisor: _____ Professor Elena Besedina

Approved by _____
Head of the KSE Defense Committee, Professor [Type surname, name]

_____

_____

_____

Date _____

ACKNOWLEDGMENTS

The author wishes to express sincere gratitude to his Thesis Advisor Dr. Elena Besedina for valuable advice, guidance and insightful comments.

I am deeply thankful to my Mother and brother Andrew for endless patience and support that cannot be overestimated.

TABLE OF CONTENTS

LIST OF FIGURES iii
LIST OF TABLES iv
LIST OF ABBREVIATIONS v
Chapter 1. Introduction 1
Chapter 2. Industry Overview and Related Studies 4
    2.1 Industry overview 4
    2.2 Related studies 7
Chapter 3. Methodology 10
    3.1 Hedonic price model 10
    3.2 Price dispersion and model specification 11
Chapter 4. Data 14
Chapter 5. Results 22
Chapter 6. Conclusions and Recommendations 27
REFERENCES 30
APPENDIX A 32

LIST OF FIGURES

Number	*Page*
Figure 2.1. Number of new publications in Ukrainian market	5
Figure 2.2 Number of published items in Ukrainian market	5
Figure 2.3 Total US Book publishers Sales Year-To-Date, Across all categories	6
Figure 2.4 – Number of Edited materials in Ukraine during 2015-2020	6
Figure 4.1 Histogram for the number of pages for Ukrainian market	15
Figure 4.2 Histogram for the price of the books for Ukrainian market	15
Figure 4.3 Scatterplot for the relationship between the price and pages for Ukrainian market	16
Figure 4.4 Books distribution by cover for Ukrainian market	16
Figure 4.5 Books distribution by Language for Ukrainian market	17
Figure 4.6 Books distribution by Genre for Ukrainian market	17
Figure 4.7 Histogram for the number of pages of the books for the US market	18
Figure 4.8 Histogram for the price of the books for the US market	18
Figure 4.9 Scatterplot for the relationship between the price and pages The for US market	19
Figure 4.10 Cover distribution for the US market	19
Figure 4.11 Genre distribution for US market	20
Figure 4.12 Histogram for different measures of price dispersion for the Ukrainian market	21
Figure 4.13 Histogram for different measures of price dispersion for the US market	21

LIST OF TABLES

Number	*Page*
Table 5.1 Estimation results for Ukrainian market	25
Table 5.2 Estimation results for the US market	26

LIST OF ABBREVIATIONS

US United States

COVID-19 Coronavirus Disease of 2019

KNCP KiltsNielsen Consumer Panel

ISBN International Standard Book Number

IDE Integrated development environment

CHAPTER 1. INTRODUCTION

Despite that the book industry is threatened by the growing digital sources of information in the United States annual sales of printed exceed 675 million .[1] E-commerce introduction and other Internet-related facilites were expected to strengthen the competition and reduce price differention (Levin, 2012).

Price dispersion stands for the differences in prices across sellers of exactly the same goods (for example, books sold at Amazon and Blackwells could be sold for different prices). Even taking into consideration distinctions in product or service quality this kind of phenomennon still holds (Ba, Stallaert and Zhang et al, 2011). Such observations also lead to the point of view that law of one price does not hold (Varian, 1980). Economic theory explains dispersion through the cost of searching. According to this explanation consumers do not have all of the necessary information about prices which let sellers to charge higher prices (Baye et al, 2005).

Since the introduction of International Standard Book Number (ISBN) selleres bacame able to computerize their catalogs with benefitting consumers and firms operating in the industry.

Even when there is complete availability of the free searching engines and websites that allow to compare prices (so addional costs of searching and investing time into it are nearly close to zero) sellers still set up different prices for the same product.

This particular type of analysis is dedicated to the online book market for economic literature and aims to understand the key drivers among different types of books.

[1] https://www.statista.com/topics/1177/book-market/#topicHeader__wrapper

Understanding the price differentials and their determinants is important for pricing strategy development and higher margins. This emprical analysis uses hedonic price model in which price dispersion is calculated for each item and is used as a dependent variable, while shipping time, availability in stock, number of pages and other attributes are treated as independent variables.

Data consists of two separate datasets for Ukrainian and US market which include prices and characterstics of books that are sold by online retailers. Data was collected using web-scraping techniques using R programming language and ParseHub application. Dataset for Ukraine is collected using the most recognizable online bookstores for the Ukrainian market (Book-ye, Yakaboo, Rozetka etc) by several foreign online bookstores. Dataset for the US includes data from Amazon, Barnes and Noble and AbeBooks.

The important implication of this work is that despite prices indeed can differ among marketplaces in Ukraine as well as in the US, price dispersion can be hardly determined by the book trim size. Instead of this we can observe for certain book genres (Fantasy and Movie Based in the US). We can also conclude that models that use the range of prices and standard deviation as a price dispersion measure give substantially different results from the dispersion measured as a skewness. Models used books length, width, number of pages, cover and genre as an independent variables.

As business recommendations we can deduce the following:

- Infrastructure and design improvement of the marketplace web-site;

- Providing the client with all of the necessari information (price, shipping costs etc.);

- Search engine optimization of the web site that will push it to the top of the search page increasing market share and number of potential buyers.

- Obtaining and monitoring information that includes peer seller prices for particulat genres, which were found to be significant contributers for price dispersion measured as a standard deviation and range of prices.

This paper is structured in the following way: the second chapter describes the recent trends and general information about online book market and publishing indistry in the US and Ukraine. The third chapter presents the basics of hedonic pricing theory and model specification. The fourth one discusses datasets and descriptive statistics for web-scraped books. The fifth chapter is dedicated to the results interpretation. The final chapter reports the most important conclusions that stem from the previous chapter and represents business recommendations.

CHAPTER 2. INDUSTRY OVERVIEW AND RELATED STUDIES

2.1 Industry overview

Books publishing industry is already a mature and well-established industry within each one homogeneous type of products is produced. Even though books can be different in their structure (for example, the same edition of the textbook can be published and sold in hardcover and paperback) and there could be some optional or complementary items sold with the book (like CD with exercises for the textbook) this type of goods is highly homogeneous.

According to the data provided at the web-site of book chamber of Ukraine (as of May 17th 2021) there was a slight decrease in both in the number of published for the first time (Figure 2.1) and number of edited materials (Figure 2.2)

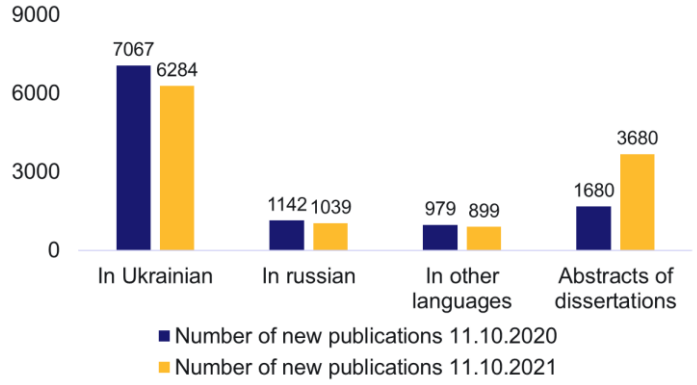
We can see that in 2021 the number of publications in Ukrainian reduced from 7067 to 6284, that amounts to more than 10% decrease, the same trend is observed for the quantity of published items in other languages, the only slight increase was in the number of publications in Russian language (from 1142 to 1039).

The same trend applies to the publications in other languages items. Number of new publications in other langages has decreased from 979 to 899, that amount to more than 10 perent decrease.

The only thing that has grown relatively to the previous year are the abstracts of dissertations. This type has grown more than 2,5 times from 1680 to 3680.

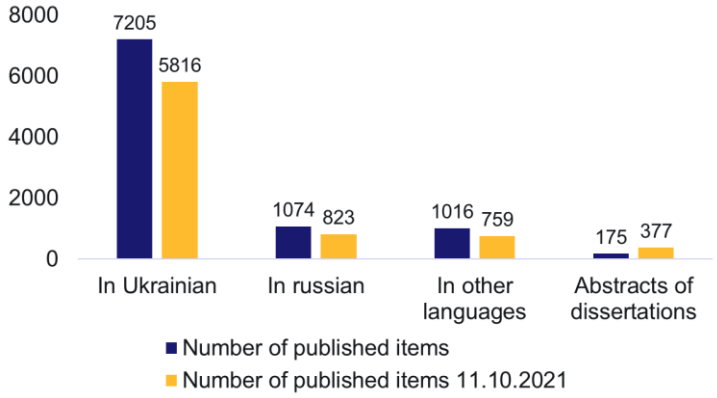
Number of the published items in 2021 also has decreased – that holds for all of the language options of the publications. For example, for materials published in Ukrainian number of printed items decreased from 1500 to 1085 thousand (by 28%) (Figure 2.2).

Figure 2.1 – Number of new publications in Ukrainian market



Source: State Scientific Institution Ivan Fedorov Book Chamber of Ukraine

Figure 2.2 – Number of published items in Ukrainian market

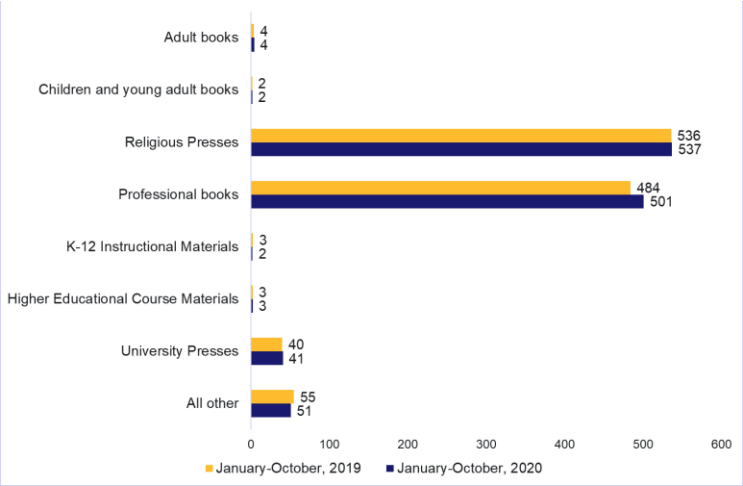


Source: State Scientific Institution Ivan Fedorov Book Chamber of Ukraine

According to the Ukrainian Association of publishers and book distributors 2019 year was one the most successful to the Ukrainian publishers, however they could not enjoy the same revenues in 20202. On the the other hand US books market was more resistant to the COVID-19 impact (Guren, Mcllroy and Sieck, 2021) (Figure 2.3, Figure 2.4).

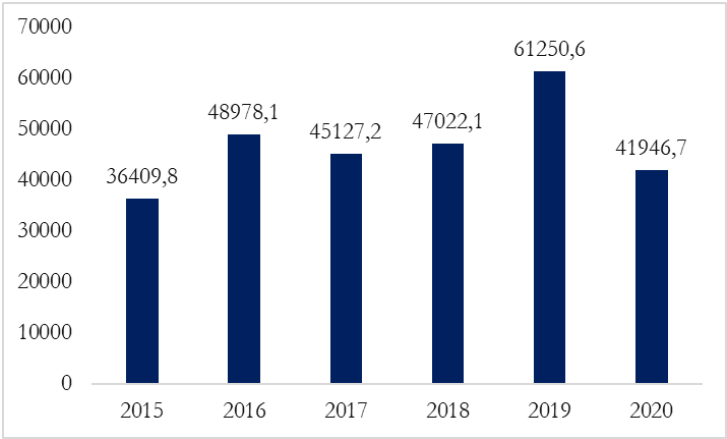
[2] http://upba.org.ua/index.php/uk/vydavnytstva/item/283-2019-naiuspishnishyi-rik-dlia-knyhovydannia

Figure 2.3 – Total US Book publishers Sales Year-To-Date, Across all categories



Source: State Scientific Institution Ivan Fedorov Book Chamber of Ukraine

Figure 2.4 – Number of Edited materials in Ukraine during 2015-2020, thousands of units



Source: State Scientific Institution Ivan Fedorov Book Chamber of Ukraine

Because of COVID-19 book industry has faced a "massives shift to online book sales" (McCarthy, Johns 2021), that is connected primarily to the quarantine restrictions adopted by many countries, since that many producers had to consider online and electronic commerce options. The same study by McCarthy and Johns also suggests that

80% of people who had an internet connection bought anything on-line during the previous month and 95% of consumers were searching for the products online.

According to the research conducted by McKinsey (Charm et al, 2020) 11% growth in the books, magazines and newspapers industry is expected. Also worth saying that people with higher income and millennials represent the leaders in "shifting spend online across both essential and nonessential items".

In general books industry is expected to enjoy growth as entire US economy, consumers spending is expected to increase which will also improve of book sales, electronic books sales growth was equal to 22%. That in its turn will make publishers to reconsider new formats of books development of "community-oriented activities" (Guren, McIlroy and Sieck, 2021). "The success of non-traditional retail outlets has highlighted that they are sometimes undervalued in sales planning".

2.2 Related studies

There are different studies that suggest using hedonic pricing to study price dispersion and its determinants as well as studies that connect market segmentation phenomenon and the books market.

For example, Luo, Zhang and Li (2018) use hedonic price function to study how factors that influence price dispersion changed over time. Their collected data for two time periods: 2001 and 2006 to compare different stages of the electronic commerce development. Sample included best sellers and random sample of books. Researchers collected prices, shipping fees, shipping time, acquisition time, dummies for availability, brand names etc.

Researchers found deviation from the "law of one price", presence of the trust assurance factor (in 2006 stores with such attribute charged higher prices), bookstore

heterogeneity could partially explain price differences, broader assortment factor became significant in the further period.

Greg Kaplan and Guido Menzio (2015) in their study on the Morphology of price dispersion found that theories of the intertemporal price discrimination were significant and studied characteristics of households that tend to pay lower prices for goods.

Findings suggest that older households that are older and have fewer employed members do so "by visiting a greater number of stores on each shopping trip instead of shopping more frequently". Their data was taken from the KNCP and included more than 300 million of transactions.

Karen Clay, Krishnan and Wolff in her study (2001) analyzed pricing of more than 30 bookstores during the period 1999-2000. She suggests that there were no changes in dispersion during the period.

Dolnicar (2008) studied market segmentation concept for tourism industry. She comes to conclusions that managers can be better off by using market segmentation and this strategy also requires product positioning.

Hedonic pricing model can be applied to analyze different markets and industries, for example Kahle and Huttel use it to estimate impact of asymmetric information in the farmland market (2020). They found out that price dispersion in farmland depends on soil quality and size of the area, also significant contribution can be brought by buyer and seller expertise. Farmers with specific knowledge and expertise face much lower search costs. So buyers that know market conditions can get lower price and sellers can find buyers with higher willingness to pay. Goldberg and Verboven apply hedonic pricing to the European Car Market (2000).

Berardi and Sevestre used hedonic to study price dispersion in retail trade (2018). According to the last mentioned study which was dedicated to grocery prices in French

supermarkets difference originated basically from permanent issues connected to the pricing in the retail chains.

To my best knoweledge, there are no studies about price dispersion drivers which analyze the Ukrainian market in general and online books market in particular

.

CHAPTER 3. METHODOLOGY

3.1 Hedonic price model

This section is dedicated to the price dispersion calculation and the theory of hedonic pricing.

In general books are goods that can be characterized by standardized and compatible characteristics. However, we should also take into consideration that process of choosing which book to purchase is quite different from the process of choosing a laptop or a smartphone.

Hedonic pricing models use the approach in which price is regressed on the attributes of the sold good which is required to be homogeneous (Koster, 2015).

The first applications of hedonic pricing models were in the agricultural and real estate markets.

According to the Koster (2015) we can also consider various forms of regression: linear, log-log or log-linear.

The hedonic pricing model can be expressed as follows:

$$Price = f(x_1, x_2, \ldots, x_n) \quad (1)$$

In this particular case $x_1, x_2, \ldots, x_n$ are characteristics of the product.

So, the price for the product would be ascertained by its qualities and utility that they can possibly bring to the potential consumer.

3.2 Price dispersion and model specification

As it was mentioned earlier price dispersion refers for the differences between prices across several different sellers of the same goods (Ba, Stallaert, Zhang, 2011). The example could be the Mankiw Macroeconomics textbook solt at Knygarnya Ye and Rozetka at different prices or carrot of the same quality, variety and origin sold at different prices in Novus and Silpo.

Following Tang et al. (2007) we will run the hedonic pricing model with the price dispersion as a dependent variable. The price dispersion itself can be calculated using different techniques. For example, Konieczny and Skrzypacz (2000) calculate price dispersion as the coefficient of variation, the same is suggested by Baye and Morgan (2005).

Other suggested measures by Berardi and Sevestre (2018) include the differences between the highest and the lowest price, interquartile range or mean absolute deviation.

As Berrardi and Sevestre in their paper calculated the corresponding measures for each of the items in the supermarket capturing the supermarkets characteristics our approach will include calculating the measure of price dispersion for each of the book with a unique ISBN.

For the completeness all of the books that are chosen for the analysis must be present at each of the marketplaces.

After the calculation of the price dispersion we will run two specifications of the model for three measures of price dispersion:

- Difference between the highest and the lowest price for the item (book with a unique ISBN);

- Variance of prices for the books sold at different market places (variance would be calculated for the books containing the same ISBN that are sold at different web sites;

- Skewness.

The following variables will be used as explanatory variables:

- Cover as a categorical variable (could be Hardcover, Paperback or other type);

- Number of pages (quantitative variable);

- Width;

- Length;

- Language as a categorical variable (Ukrainian, Russian, English or Other);

- Genre as a categorical variable (Classic, Movie Based, Comics, Child, Business, Science Fiction, Foreign Literature, Psychology, Ukrainian literature, Fantasy, Textbooks or Vocabularies).

Thus, our task would be to estimate and interpret 3 following models:

$$Range\ in\ prices = Coverage + Number\ of\ pages + Length + Width + \ anguage + Genre + Cover; \quad (2)$$

$$SD\ of\ Prices = Coverage + Number\ of\ pages + Length + Width + anguage + Genre + Cover; \quad (3)$$

$$Skewness = Coverage + Number\ of\ pages + Length + Width + Language + Genre + Cover. \quad (4)$$

Skewness would be calculated using the following formula (as it provided for the respective function in the R programming language:

$$Skewness = 3 * \frac{Mean - Mode}{Standard\ deviation} \quad (5)$$

For each of the unique items (book that have the same ISBN) we will calculate Skewness as well as other measures as dependent variable and plug into regression as dependent variable.

CHAPTER 4. DATA

Data was web-scraped using ParseHub application and R programming language. Using two different tools was necessary to collect as many as possible data and due to computational issues. There is no possibility to run at least two R data web scraping as well as 2 ParseHub scraping sessions simultaneously.

The bigger amount of data was dictated not only by needs of convincing and reliable regression results but also the circumstances due to which it was required to have at least the price of the item at each of the marketplaces. If there is no price for all of the marketplaces of a single unique item (unique book) we had to exclude this price from the analysis, as the range of the prices and variance of the prices (dispersion measure other words) would not be comparable.

For Ukrainian market data was scraped from 4 Ukrainian online marketplaces which included: "Rozetka", "Knygarnya Ye", "Yakaboo" and "Bukva Ua".

The data was collected obtaining the URL for each of the book, with the following passing to the page itself scraping the predefined features: price, pages, cover, size characteristics etc. After the data from the first page was scraped pagination passing methods were applied in order to collect data from the multiple pages.

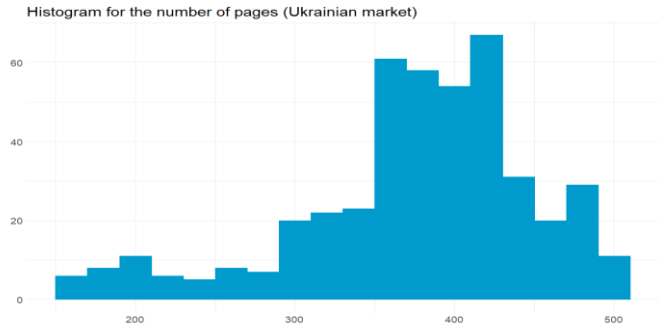
After the data from different web-sites and pages was scraped it was merged, cleaned and rearranged using Excel and R programming language with R studio IDE.

If the book had complete observations with price and ISBN at all of the web sites it was included to the dataset. There was a part of observations which had missing explanatory characteristics items but had price and ISBN presented at all of the market places, partially such observations were completed using the information from other available web sites.

The resulting data set for Ukrainian market consists of 427 observations.

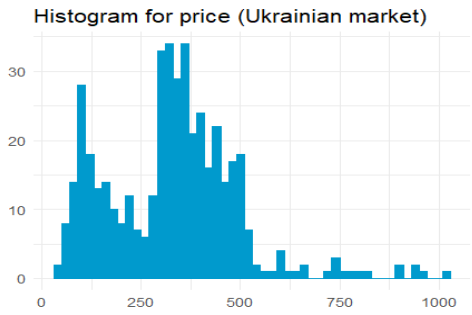
Most of the books in the collected dataset has the mean number of pages that is equal to 375, mean price for the book is equal to the 333 UAH and the median price for the book is 340 UAH.

Figure 4.1. Histogram for the number of pages for Ukrainian market



As we can see most of the books have price that is from 300 UAH to 500 UAH and there is only a few books which price is more than 500 UAH (visuals with prices were created using the closest the prices that are presented at Rozetka retailer).

Figure 4.2. Histogram for the price of the books for Ukrainian market

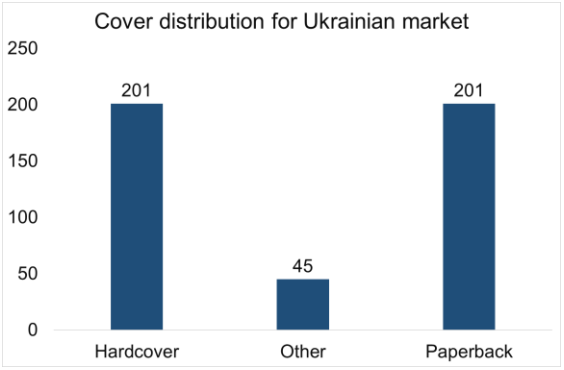


Figures 4.3 describes the relationship between the price and number of pages in the book. As we can see there is no straightforward relationship between the price and number of pages for Ukrainian market.

Figure 4.3. Scatterplot for the relationship between the price and pages for Ukrainian market

40% of books had a Hardcover cover and 40% of books were Paperback, the rest of the books has other type of cover (see Figure 4.4).

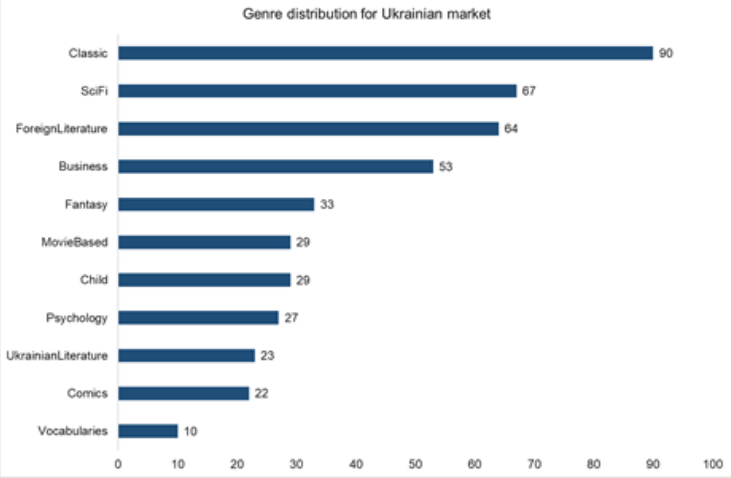
Figure 4.4. Books distribution by cover for Ukrainian market



Most of the books were written in Ukrainian 297 (or 66%) and most of the books were Classic (65 or 18%), Science Fiction (67 or 14%) and Foreign literature (64 or 13%). Results are presented on the Figures 4.5 and 4.6.

Figure 4.5. Books distribution by Language for Ukrainian market

Figure 4.6. Books distribution by Genre for Ukrainian market



For the US market most of the books (nearly 60%) have number of pages that is less than 350, there are also not so many books for which number of pages is more than five hundred (Figure 4.7).

Pricing distribution for US market is also quite similar to Ukrainian with only difference in magnitude on the horizontal axis which is caused by different currencies.

Figure 4.7. Histogram for the number of pages of the books for the US market

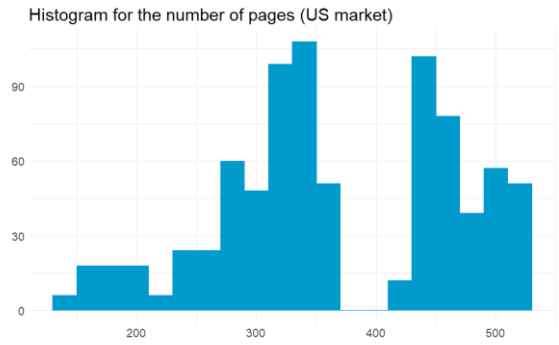
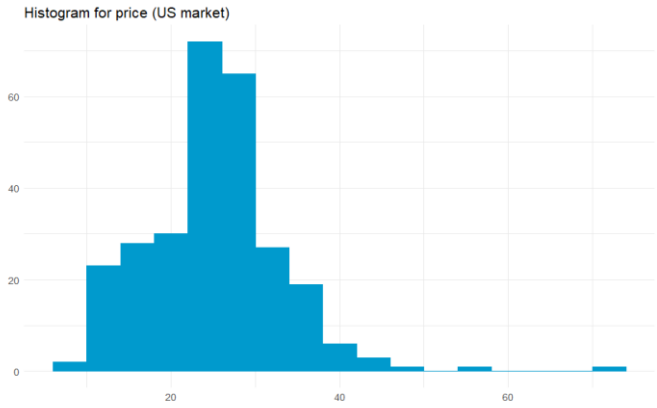
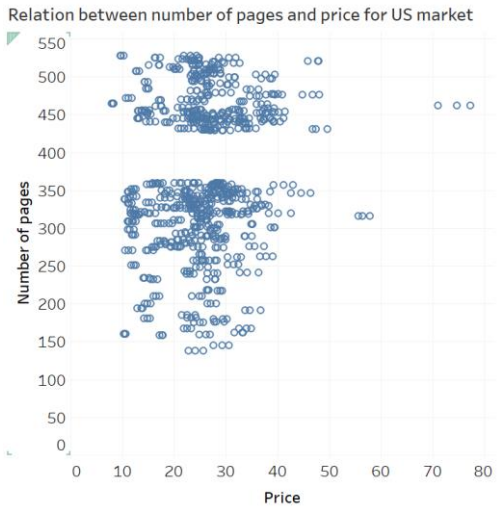


Figure 4.8. Histogram for the price of the books for the US market



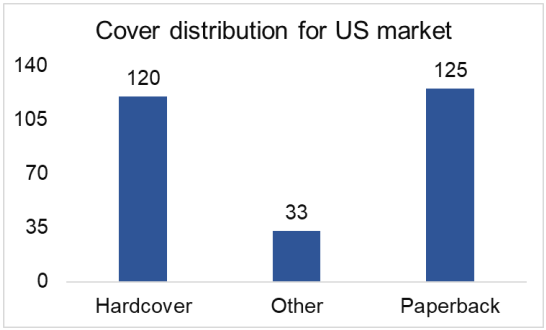
As well as in the Ukrainian case there is no clear and evident dependence between the price and number of pages for the US market (Figure 4.9). Most of the web-scraped books belong to Science fiction and Fantasy genre (89 and 97 respectively).

Figure 4.9. Scatterplot for the relationship between the price and pages for US market



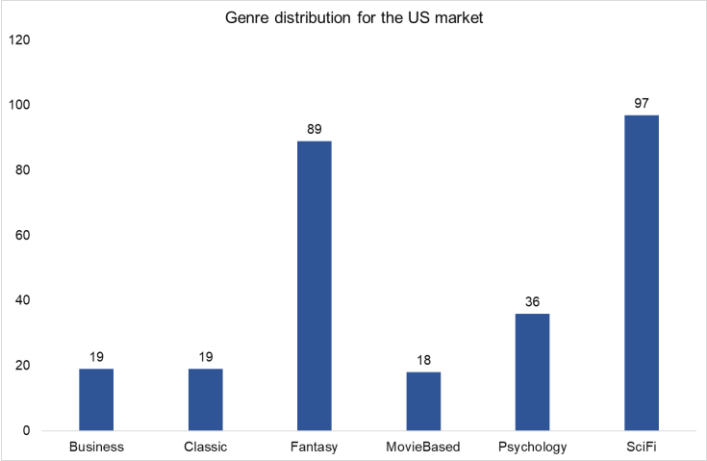
The biggest share of the items that have been scraped for the US market were paperback. The number of paperback items equals to 125 which accounts to 45% of the entire dataset, there are 43% (or 120 in absolute terms). The rest of the books are made with other type of cover (Figure 4.10).

Figure 4.10. Cover distribution for US market



Among the books that were included into the US dataset most of the books are Science fiction books, there are 34% or 97 of such books. The second biggest group includes fantasy books, number of such books amounts to 89 items in absolute terms or 32%. The rest 92 books or 34% is distributed among business, comics, movie based and psychology books.

Figure 4.11. Genre distribution for the US market



As it has been expected different measures of price dispersion have different distribution (Figure 4.12, 4.13). Skewness distribution is the most left skewed among others having the lowest mean value which is equal to zero for US market, at the same time mean value for the range in prices is equal to 2.41 and for standard deviation measure to 1.25 (Appendix A).

For the estimation results for Ukrainian market the same situation is observed: mean value for range is equal to 35 and the mean for skewness is 0. Mean of standard deviations is 15.

Given the information above we can decide that among those 3 measures the most appropriate one is standard deviation, as it captures prices from all of the retailers and has an appropriate interpretability.

Figure 4.12. Histogram for different measures of price dispersion for the UKR market

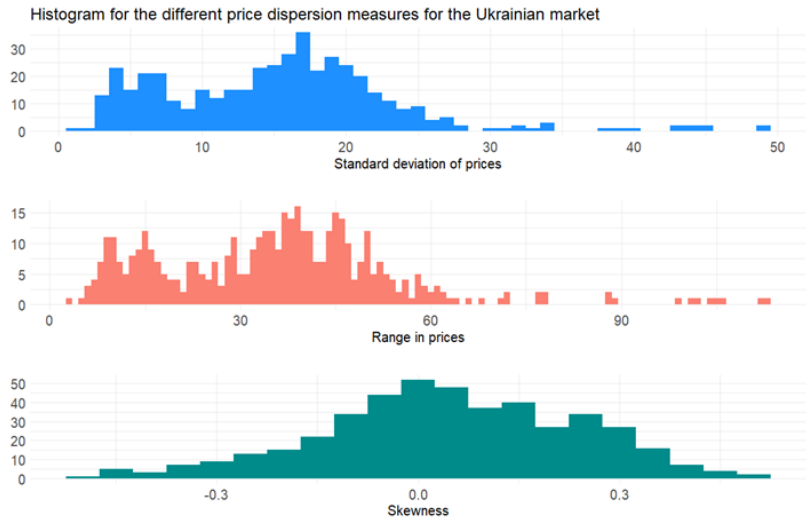
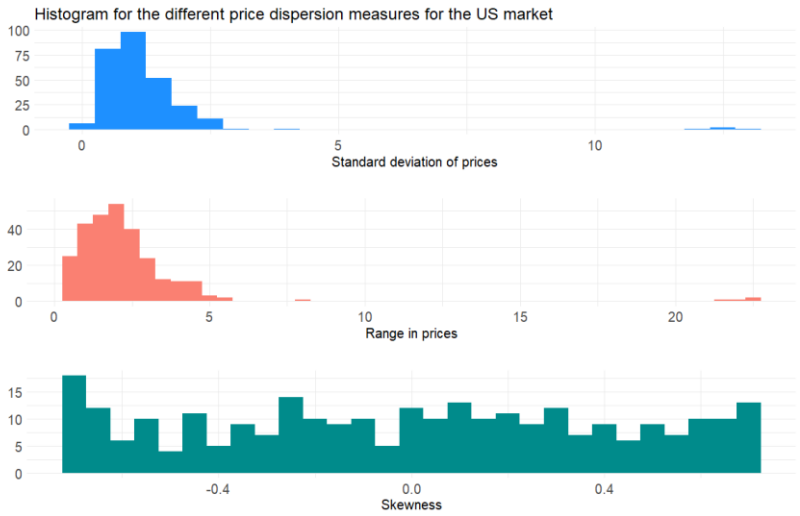


Figure 4.13. Histogram for different measures of price dispersion for the US



CHAPTER 5. RESULTS

This chapter is dedicated to the estimation results of the price dispersion for the online book market.

For Ukrainian market 3 models have been estimated. The first model corresponds to the case when price dispersion is measured as a range of prices, the second one as the variance and the third one as a skewness.

Models have shown similar results in terms of significant coefficients, however for some cases range of prices and standard deviation have shown different signs of coefficients (Table 5.1).

We can conclude that despite that prices can differ among retailers (web-sites) for both Ukrainian and the US market, price dispersion phenomenon can be barely explained by books characteristics. Based on our estimation we can judge only about contribution that can be brought by some of the book characteristics (for example, genre or some distinct type of cover).

We can also conclude that according to our research using skewness as a dependent variable gives us nearly no explanatory information due to the extremely low R-squared an absence of the overall regression significance (Table 5.1).

Coefficients near variables that indicate whether book belongs to comics, foreign literature, psychology and science fiction genre are statistically significant for the range of prices for Ukrainian market dependent variable.

Mentioned above dummy variables have coefficients with positive signs. For example, if book genre is comics, so price dispersion measured as a range of prices within this genre would be on average by almost 11 hryvnias higher, relative to the business

literature which is a base category, this conclusion is made using the 0.05 significance level.

If book belongs to science fiction or psychology genre price dispersion measured as a range of prices for these measures would be by 22 and 20 hryvnias higher respectively (relative to the business literature). However, the results are statistically significant only at the 0.1 significance level.

There are also genres that appear to be significant with negative signs, which implies that range of prices decreases with books affiliation to the corresponding genre. This effect is present for classic books, Ukrainian literature and vocabularies.

According to the model that uses the standard deviation as a dependent variable, books in psychology and foreign literature genres tend to have higher price dispersion by 7.2 and 9.6 hryvnias relative to the business literature.

It is worth mentioning that there are different directions (signs) of coefficients for fantasy and science fiction genres for Ukrainian market. This can be explained by some unobserved effects (such as temporary discounts that appear on retailer web-sites etc.) or by construction of a measure as range captures only maximum and minimum value but ignores the placing and magnitude of two other observations, unlike standard deviation measure that involves mean value and values of all the observations for calculation.

Number of pages, width and length turned out to be insignificant for all of the models for Ukrainian market.

For the US market data, we estimate models that use range and standard deviation as dependent variables. Results are presented in Table 5.2.

Number of pages seem to be to be significant only for the US market. However, the absolute value of the coefficient is very low for both markets (US and Ukrainian markets). That implies that number of pages could not be the important determinant of price dispersion due to its low magnitude.

The same logic applies to the width and length coefficients. Coefficients are insignificant for all of the measures and both of the markets.

The one of the possible explanations could be that retailers do not position hardcover or some specific types of books to be more luxurious or premium products apart from paperback, despite that obviously working with hardcover is more costly as it involves more paper and other materials to produce, hardcover books also weight and take more space while transporting so it can lead be additional costs.

Other types of cover have been found to be significant determinant of price dispersion for the US market.

For dispersion measured as a range of prices having other type of cover increases price dispersion relative to the base category (hardcover) by 1.2 dollars and accordingly by 0.6 US dollars for price dispersion measured as a standard deviation of prices.

According to the estimation results, we can also conclude that for Ukrainian and US markets the most significant coefficients appear near genre variables. It is also worth mentioning that they can be both positive and negative and given that genre variables are dummy variables we can infer that there are differences in price dispersion that are caused by genre.

For the US market coefficient near Movie based category was found to be significant. That implies that relative to the business books range in prices is on average lower by 1.8 US dollars and the standard deviation is lower by 1 US dollars.

Table 5.1. Estimation results for Ukrainian market

Dependent variable:	Range of prices		Standard deviation of prices		Skewness of price distribution	
	Coefficient	SE	Coefficient	SE	Coefficient	SE
Length	-0.082	0.067	-0.037	0.029	0.0003	0.001
Width	0.078	0.052	0.035	0.022	-0.0004	0.001
Number of pages	0.006	0.011	0.003	0.005	-0.0001	0.0001
Language: Russian	0.874	4.376	0.263	1.894	0.065	0.049
Language: Ukrainian	3.716	4.628	1.476	2.003	0.058	0.052
Genre: Child	-1.678	3.958	-0.816	1.713	0.003	0.044
Genre: Classic	-7.389**	3.034	-0.816	1.313	-0.017	0.034
Genre: Comics	10.997**	4.377	-3.241**	1.894	0.005	0.049
Genre: Fantasy	-8.664*	4.851	-3.826*	2.100	0.014	0.054
Genre: Foreign Literature	16.632*	3.347	7.256***	1.449	-0.007	0.038
Genre: Movie Based	-17.092**	4.059	-7.543***	1.757	0.057	0.045
Genre: Psychology	22.039*	4.118	9.635***	1.783	-0.037	0.046
Genre: Science Fiction	20.298*	3.184	-8.890***	1.378	-0.017	0.036
Genre: Ukrainian literature	-24.850***	4.337	-10.903***	1.877	0.041	0.049
Genre: Vocabularies	-28.152***	6.037	-12.390***	2.613	0.040	0.068
Cover: Other	0.547	2.956	0.231	1.279	-0.014	0.033
Cover: Paperback	1.802	1.804	0.787	0.781	-0.008	0.020
Constant	38.183***	7.584	16.602***	3.283	0.064	0.085
Observations	447		447		447	
R2	0.192		0.195		0.022	
Adjusted R2	0.160		0.163		-0.016	
Residual Std. Error	17.072		7.390		0.191	
F Statistic	6.007***		6.128***		0.574	

Note: Stars denote level of significance with *p<0.1; **p<0.05; ***p<0.01

Table 5.2. Estimation results for US market

Dependent variable:	Range of prices		Standard deviation of prices	
	Coefficient	SE	Coefficient	SE
Length	0.295	0.284	0.143	0.158
Width	-0.232	0.308	-0.102	0.172
Number of pages	-0.005***	0.002	-0.003***	0.001
Genre: Classic	-1.193	0.929	-0.672	0.518
Genre: Fantasy	-1.200*	0.720	-0.701*	0.402
Genre: Movie Based	-1.768**	0.893	-1.010**	0.498
Genre: Psychology	-0.099	0.788	-0.144	0.439
Genre: Science Fiction	-0.631	0.704	-0.357	0.393
Cover: Other	1.207**	0.486	0.685**	0.271
Cover: Paperback	0.285	0.332	0.160	0.185
Constant	4.914**	2.022	2.561**	1.127
R2	0.072		0.071	
Adjusted R2	0.044		0.043	
Residual Std. Error	2.801		1.562	
F Statistic	2.597***		2.548***	

Note: Stars denote level of significance with *p<0.1; **p<0.05; ***p<0.01

CHAPTER 6. CONCLUSIONS AND RECOMMENDATIONS

This term-paper provides an outlook and interpretation of price-dispersion determinants basing on 2 models and web-scraped data for the US and Ukrainian market.

In this work I have used 3 types of price dispersion estimation models that include the range between maximum and minimum price, standard deviation and skewness. It is found that models provide different results. However, for some cases significant coefficients overlap between model and are quite close in terms of their absolute values.

We can conclude that despite that prices can differ among retailers for both of Ukrainian and US market, price dispersion phenomenon can be barely explained by books characteristics such as number of pages, length and width. Price dispersion can be increased basing on the fact that book belongs to the certain genre or has some specific type of cover. According to the estimations we have found out that number of pages was found to be significant only for the US market, but value of the coefficient leads us to the conclusion that number of pages could not be the determinant of price dispersion due to its low magnitude. Width and length coefficients are insignificant for all of the measures and both of the markets. For dispersion measured as a range of prices having other of cover increases price dispersion relative to the hardcover) by 1.2 dollars and accordingly by 0.69 for price dispersion measured as a standard deviation of prices.

According to the estimation results, we can also conclude that for Ukrainian and US markets, the most significant coefficients appear near genre variables. It is also worth mentioning that they could be positive and negative. Since that we can infer that if book belongs to particular genre that could be a sign that prices of the different retailers can be closer to the average value – that is observed for Ukrainian literature books, Vocabularies and Movie Based books. We can also see the similar situation for books sold at the US market that belong to Fantasy and Movie Based genres.

The absence of the overall explanation of variation in models can be explained by that during the recent 15 years searching costs has decreased a lot, search engine developer companies invest into research and development and also permanently gather data about user which lets to optimize the search. Companies (book retailers in this case) also invest into search engine optimization departments and services. Since that match between the searched product, buyer and seller occurs much faster. For example, WebFX[3] company provides SEO not only for Google search. Another reason could be that buyers do not require specific expertise unlike from farmland, privatization object, or used car. Buyers choose books basing on their reading experience, needs for education or other unobservable factors. The possible extension of the work can include the influence of search engine optimization investment on the searching costs and time and other factors mentioned above.

Based on the research we can derive the following recommendations. First of all, due to the fact that only certain distinct book characteristics contribute into price dispersion we can infer that online retailers should care about improvements of the web-site or application infrastructure in order to deal with consumers that are willing to pay certain price exactly within their marketplace. This can include special offers, discounts and collecting personalized information about clients. It is also worth mentioning that retailers should not overindulgence with client information possession as being involved into scandals with leakage of personal information would decrease trustworthiness.

The second recommendation connected to the client engagement includes providing the client with all of the information about the good, for this case it may include reviews, description, price clarification shipping availability etc.

[3] https://www.webfx.com/amazon/amazon-seo-pricing.html

The third recommendation is related to the search engine optimization of the web site. The consequence would be that web site would be closer to the top on the search page. This would lead to higher reliance, market share and number of potential visitors.

The fourth recommendation involves monitoring prices of the books in particular genres (where price dispersion was found to be greater than other categories) and aggregation of the peer seller prices. This will let retailers have more perfect information about market conditions and demand, since menu costs in e-commerce are very low that can be seen as a considerable way for sales growth and margins improvement.

REFERENCES

Baye, et al. 2005. Information, Search, and Price Dispersion, *Handbook on Economics and Information Systems (Elsevier, T. Hendershott, ed.).*

Berardi et al. 2017. The Determinants of Consumer Price Dispersion: Evidence from French Supermarkets. *Working papers 632, Banque de France.* https://ideas.repec.org/p/bfr/banfra/632.html

Clay Karen et al. 2001. Prices and Price Dispersion on the Web: Evidence from the Online Book Industry. *Journal of Industrial Economics* 49 (August): 521-39.

Goldberg Pinelopi, and Verboven Frank. 2001. The Evolution of Price Dispersion in the European Car Market. *The Review of Economic Studies* 68 (October): 811-848.

Guren, et al. 2021. COVID-19 and Book Publishing: Impacts and Insights for 2021. Published on 19th February 20. https://www.scienceopen.com/document?vid=79fd33e1-5f35-4414-9b99-16f64da78305

Hal Ronald Varian. 1980. A Model of Sales. *The American Economic Review* 4 (September): 651-659.

Hans Koster, 2015. Lecture slides on Hedonic Pricing: https://www.urbaneconomics.nl/wp-content/uploads/2015/09/Hedonic-pricing.pdf

Jifeng Luo, Han Zhang, and Haizheng Li. 2017. Pricing strategies in online book industry: a comparative study. *Information Systems and e-Business Management* 4 (June)*:* 791-816.

Kaplan Greg, and Menzio Guido. 2015. The Morphology of Price Dispersion. *International Economic Review* 56 (October): 1165-1206.

Konieczny Jerzy, and Skrzypacz Andrzej. 2000. The behavior of price dispersion in a natural experiment. *Stanford university.*

Levin Jonathan. 2012. The Economics of Internet Markets. *Stanford University and NBER.* https://web.stanford.edu/~jdlevin/Papers/InternetMarkets.pdf.

McCarthy Pete, and Johns Jess. 2021. COVID-19 and The Publishing Industry, Part 1: A Year in Review, Consumer Trends & Behaviors. Published on March.

Sara Dolnicar. 2008. Market segmentation in tourism. *Tourism Management: Analysis, Behaviour and Strategy* (December): 129-150.

Seifert Stefan, Kahle Christoph, and Hüttel Silke. 2020. Price Dispersion in Farmland Markets: What Is the Role of Asymmetric Information. *American Journal of Agricultural* 103 (October): 1545-1568.

Sulin Ba, Jan Stallaert, and Zhongju Zhang. 2011. Research Note - Online Price Dispersion: A Game-Theoretic Perspective and Empirical Evidence. *Information Systems Research* 23 (2): 575-592.

Tamara Charm et al. The great consumer shift: Ten charts that show how US shopping behavior is changing. Published on 4th August 2020.

Tang Zhulei et al. 2010. The Impact of Shopbot Use on Prices and Price Dispersion: Evidence from Online Book Retailing. International Journal of Industrial Organization 28(6): 579-590.

The New Palgrave Dictionary of Economics, Published on December 3, 2016 https://link.springer.com/referenceworkentry/10.1057%2F978-1-349-95121-5_1976-1.

APPENDIX A

Table 1. Descriptive statistics for different measures of price dispersion

Market	Measure	Min	First quartile	Median	Mean	Third quartile	Max	Standard deviation	Skewness
US	Range in prices	0.290	1.273	1.995	2.408	2.748	22.700	2.659	6.033
US	Skewness	-0.707	-0.343	0.007	-0.006	0.326	0.707	0.429	-0.027
US	Standard deviation	0.146	0.680	1.022	1.253	1.409	12.859	1.484	6.403
UKR	Range in prices	3.270	21.945	36.320	35.530	45.515	113.080	18.630	0.958
UKR	Skewness	-0.489	-0.072	0.045	0.048	0.184	0.507	0.190	-0.211
UKR	Standard deviation	1.400	9.385	15.753	15.373	19.591	49.122	8.080	0.951