Price strategies of the international airline market.

Alexandra Belova

March 1, 2015

WORKING PAPER. PRELIMINARY

Abstract

The focus of this paper is on the high dispersion of the airline ticket prices which appeared with a new pricing policy, called yield management. The paper empirically investigates the connection between ticket prices on the international market, characteristics of the flight, consumer personal characteristics and their willingness to pay. The main contribution of the paper is connected with the employment of a new database: Survey of International Air Travelers (2005). This data base allows us to control for the parameters which can have crucial influence on the prices, like reservation moment and personal characteristics of consumers. The moment of reservation plays a leading role in our analysis as it is the main instrument of the firm's price strategies.

JEL Classification: L11, L93, D49

University Paris 1 Pantheon Sorbonne - Paris School of Economics
Adresse: 106-112 Boulevard de l'Hopital, 75013, Paris, France, Bureau 223
Tel: +33 7 50 36 74 94
Email: alex.v.belova@gmail.com; alexandra.belova@univ-paris1.fr

1 Introduction

Few decades ago the world market of the passenger airline service was deregulated.[1] A lot of articles concerning the competition analysis of the deregulated market have appeared since liberalization. The majority of them considered the airline market homogeneous, which means all the tickets for the same route are assumed to have the same price.[2] However, several years after the liberalization the ticket prices became highly diverse on any given route, so the assumption of homogeneity is not always justified. In the current paper we try a new approach and construct a pricing model in order to explain such a price diversity.

Fierce competition caused by liberalization made the companies introduce a new marketing approach based on consumers' preferences, named yield management.[3] The idea of yield management is to differentiate the airline service products in order to have the possibility to charge different types of consumers with different prices. The development and use of yield management models have resulted in companies offering a variety of different types of fares for the tickets on the same route. Its main instruments: ticket class, service restriction and reservation moment - allow airlines to divide consumers into groups according to their elasticities.[4] Therefore price diversity has a direct connection with the consumer preferences. That is why in our model we introduce consumer characteristics. The main goal of this paper is to estimate a relevant pricing model taking into consideration principal instruments of the yield management and clients behavior. We analyze the connection between prices, characteristics of the flight and consumer personal attributes.

For any kind of price model it is necessary to control the huge price dispersion and product differentiation introduced by the yield management policies. Most of the first articles devoted

[1] Originally the whole air service market was strictly regulated by the range of government agreements. In the 1970s the U.S. and later in the 1980s Europe deregulated the air service market enabling companies to choose their own prices, routes and volumes of production and allowing new companies to enter the market. Transatlantic market was deregulated in 1990s.

[2] Barley, Pazar (1981), Strassmann (1990), Dresner, Tretheway (1992), Brueckner, Dyer, Spiller (1992), Brueckner, Spiller (1994), Marin (1995).

[3] The expression "yield management" was invented by Robert Crandall, former Chairman and CEO of American Airlines, who has called it "the single most important technical development in transportation management since we entered deregulation".

[4] "A reservation moment" - how many days prior to departure a ticket was bought

to airline competition analysis consider the airline market to be homogeneous.[5] Later in a number of articles the authors pay attention to a huge dispersion in prices on the same route. Borenstein (1989), Dresner and Tretheway (1992), Evans and Kessides (1993) try to control for several level of prices in their research.[6] A paper of Borenstein and Rose (1994) is the first one with the main focus on price dispersion. They prove that the pattern of the observed price dispersion cannot be explained by cost differences alone. They argue that the data support the models of price discrimination and also agree that they do not have sufficient data to test all the theories and make their analysis exhaustive.[7] There are also a few following papers analyzing the connection between competition and price dispersion on the air service market.[8]

Unlike Borenstein and Rose (1994) we explain high dispersion by the consequences of the yield management approach. There are two possible reasons, why this topic has not received enough attention in the literature. Firstly, it is not easy to find a reliable data with detailed information about flights and consumers like reservation moment or personal characteristics. Secondly, in order to employ a model of differentiated market we consider the reservation moment being a product characteristic. The product stays the same regardless how many days prior to the departure it was bought, but companies discriminate consumers by the moment of purchase. We assume that consumers are aware of companies' policies, so in equilibrium maximizing their utility they choose the best moment of purchase as another product characteristic.

The principal contribution of the current article is connected with a new database, which has not been employed before for the competition analysis of the airline market. It is the Survey of International Air Travelers (2005) database, made by American International Trade Administration Office of Travel and Tourism Industries. This survey was conducted in the

[5] Barley, Pazar (1981), Strassmann (1990), Dresner, Tretheway (1992), Brueckner, Dyer, Spiller (1992), Brueckner, Spiller (1994), Marin (1995).

[6] However, the main goal of Borenstein (1989) was to analysis influence of market share and airport share on prices, among the conclusions we see that low-end fares appears to be more responsive to the competition. Similar conclusions are made in articles of Dresner and Tretheway (1992) and Evans and Kessides (1993): deregulation has reduced the discounted prices but have not influenced the highest fares on the market. All these articles fail to explain how this price discrimination in possible, as they do not have access to data which allows distinguishing the consumers types.

[7] Destination and Origin Survey (DB1B, by Department of Transportation) is the database employed by Borenstein and Rose (1994)

[8] Stavins (2001), Clemons and all (2002), Giaume and Guillou (2004), Kristopher and Shapiro (2009)

airports of the U.S. and it involved a long list of questions about the flight, prices, activities during the trip, personal characteristics of the polled and their opinion about quality of the services. The main advantage of the Survey of International Air Travelers (2005) is that it provides the consumers' characteristics and all the details about the sold tickets. It gives us possibility to control for ticket class and date of departure. The travelers were also inquired how many days prior to the departure the ticket had been bought. That allows distinguishing between the consumers who buy tickets earlier and closer to the departure date. The databases, similar to the Survey of International Air Travelers, are usually employed for the travel demand studies, like Reece (2001), Divisekera (2003), Racello and all (2005), Laesser and Crouch (2006), Cortes-Jimenez and Blake (2010).[9]

The current article analyzes the market of the airline services from a totally new perspective. We conduct a price analysis controlling for the main tools of the yield management. For this purpose we employ a hedonic approach which is specially developed for the markets with product differentiation.[10] We construct a model in order to study the connection between product attributes, prices and consumers characteristics. To achieve this goal we estimate a system of equations. The first one is a hedonic price function: price as a function of the product characteristics. The others are equations of the consumers' willingness to pay (WTP) for one or another product characteristic depending on the consumers' individual attributes like income, sex or age.

The rest of the paper is organized as follows. The next section presents a few details about the yield management. Section 3 is devoted to the model. Section 4 provides a detailed data description and descriptive statistics of the variables. The empirical results are described in Section 5. Section 6 concludes.

[9]Reece (2001) analyzes how the demographic, lifestyle and distance variables explain household travel to Los Vegas and Atlantic Cape May. Divisekera (2003) study the substition effects between different destination of the US tourists. Racello and all (2005) develop an dynamic model of demand employing the information variable. Laesser and Crouch (2006) estimate also a hedonic price function for overall expenditures of the travelers. Cortes-Jimenez and Blake (2010) use a country level aggregate data to construct a dymanic model of touristic demand.

[10]For details see Rosen (1974), Nesheim (2006)

2 Yield management

In different industries companies use a great number of approaches in order to understand the consumers behavior and sell their products with the highest profit margin. One of them is "yield management", a set of special pricing strategies developed by airline companies after the deregulation of the market. This strategy let airline companies survive without government support. For the last few decades the airline companies all over the world have started to use yield management. Yield management instruments are also employed by a number of other industries, e.g. hotels, transportation, telecommunication, insurance, rental, advertising companies, etc.

Yield management is the process of understanding, anticipating and influencing the consumer behavior in order to maximize profits (or yield) from a fixed, perishable resource (such as an airline seat on a plane or a hotel room or an advertising space). So the product can be bought in advance for some particular moment, and if it is not sold, it cannot be stored. By another definition yield management involves strategic control of the available resources, meaning to sell the product to "the right customer at the right time for the right price".[11]

The general idea of the yield management is to divide the consumers into the groups according to their elasticities and charge every group with a different price. There is a number of ways to identify the groups of consumers: business and economy classes, service restrictions, loyalty program, etc.[12] A consumer choosing the services rendered to her provides a company with the information about her preferences that helps to define which group this consumer belongs to. For examples, business class travelers are less price sensitive than the economy class passengers. But even within the single service compartment case, for examples, the economy cabin on an airplane, the companies sells the products for different fares. In this case they employ intertemporal price discrimination, which means the consumers are offered different prices if

[11] American Airlines annual report, 1987, definition of the yield management

[12] see Smith and all (1992), Weather and Bodily (1992), Botimer and Belobaba (1999), Coughlan (1999), Desiraju and Shugan (1999), Feng, Xiao (2000), Chen and all (2003)

they purchase the products further or closer to the departure date.[13]

The companies change the prices during the whole period of sales as the different type of consumers prefer to buy or book the products at different moments.[14] For example, in case of airline or hotel industry early arrivals are generally individuals or families who are planning vacations or other trips far in advance. These customers are usually price sensitive and have lower costs of committing to an early purchase. On the other hand for business or other urgent trips the consumers have to book their tickets or room closer to the date of departure because they did not plan their travel earlier. In general the consumers who buy product at the last moment are characterized by a lower price elasticity of demand. In a number of yield management models, described in the literature, the demand for each fare product is treated even as completely separable and independent.[15] Although these assumptions are clearly unrealistic, the products cannot be considered being absolutely homogeneous either. Despite the fact that in the literature it is called "intertemporal price discrimination" yield management is not an example of price discrimination.[16] In case of price discrimination companies also use differences in consumer elasticities, but they are not obliged to sell the fixed number of tickets before departure.[17]

The total number of economic papers devoted to the yield management strategy is very small. Moreover the majority of them presents either qualitative analysis (Smith and all (1992), Weatherford and Bodily (1992), Weatherford and Belobaba (2002)) or develop theoretical framework (Botimer and Belobaba (1999), Coughlan (1999), Dana (1999), Desiraju and Shugan (1999), Feng and Xiao (2000), Chen and all (2003)).[18]

[13]In the most of the articles, focused on yield management, the differentiation of the prices in time is mentioned as the one of the principal tools of the yield management: Smith and all (1992), Weather and Bodily (1992), Dana (1999), Desiraju and Shugan (1999), Feng and Xiao (2000), Chen and all (2003)

[14]Borenstein and Rose (1994) called it self-selective price discrimination: "self-selective price discrimination relies on product heterogeneity, since it is carried out by offering consumers a set of alternatives and allowing their choices to reveal information about their characteristics".

[15]This kind of approach can be found in the following articles: Botimer and Belobaba (1999), Coughlan (1999), Dana (1999), Feng and Xiao (2000), Chen and all (2003).

[16] Borenstein and Rose (1994).

[17]Dana (1999) devotes an article to this topic showing the difference between price discrimination and yield management.

[18]In some of these articles there are several simulations but there is no real data estimations.

3 Model

3.1 Hedonic price function

For the purpose of this paper we use hedonic approach which is specifically developed for markets with a differentiated products. Firstly introduced by Rosen (1974), a hedonic price function describes the equilibrium relationship between the economically relevant characteristics of a product (or service) and its price. It is used to measure consumer and producer valuations of differentiated product and to calculate quality adjusted price indexes for goods. Rosen describes a model of product differentiation based on the hedonic hypothesis that goods are valued for their utility-bearing attributes or characteristics. According to hedonic approach any class of the differentiated products can be completely described by a vector of measurable characteristics. Observed product prices and product characteristics define a set of implicit (or "hedonic") prices which guides both consumers and producers to make their decision in the characteristics space. According to their characteristics different companies choose to produce goods with some particular set of attributes. For examples, in case of the airline market the low-cost companies tend to provide only products with a minimum number of services, while the huge well-established companies find more profitable to sell the tickets with a full range of services. At the same time different types of consumers choose different products according to their preferences. Consumers maximize their utility and companies maximize their profit choosing the best set of product attributes and prices. Competition prevails because a single agent adds zero weight to the market. The equilibrium prices and the functional forms fundamentally are determined by the distribution of buyers and their preferences, the distribution of sellers and their costs, and structure of competition in the market.

3.2 First step. Price function estimation

This chapter is devoted to the model specification chosen for the empirical estimation. The following notation will be used in the model.:

$j = 1...J$ - product index,
$i = 1...I$ - consumers index,
$t = 1...T$ - market (route) index,
$l = 1...L$- airline company index,
$p_{j,t,l}$ - product price,
x_j - vector of the product characteristics
$k = 1...K$ - product characteristic index (K is the size of the vector x).

As it is recommended by Rosen, as the first step the standard hedonic price function should be estimated by a standard hedonic method regardless the firm and consumers characteristics:

$$p_{j,t,l} = f(x_j, \delta_t, \alpha_l) \tag{1}$$

where α_l and δ_t are the company and route effects. These fixed effects are also considered being product characteristics, as they are observable to the consumers and they have an influence on their choice. The specific fixed effects are incorporated to account for unobservable characteristics of the route and company.

The main problem of hedonic price that there is no theoretical specification for the price function (1). A number of researches argue that hedonic price function should be estimated nonparametrically or semiparametrically.[19] But in this article we do not apply the nonparametric estimation. Firstly, in case of nonparametric function we do not have possibility to interpret coefficients. Secondly, there are not enough observations for any kind of the nonparametric estimation.

For the estimation of the hedonic price function we chose log-linear functional form:[20]

$$log(p_{j,t,l}) = \sum_k log(x_{j,k})\gamma_k + \delta_t + \alpha_l + \epsilon_{j,t,l} \tag{2}$$

These is a general form of the estimated functions. Following Pakes (2003), we are going to estimate several alternative specification of price function in order to find the best specification.

[19]Bajari, Benkard (2001), Ekeland, Heckman, Neisheim (2004), Neisheim (2006)
[20]see Coulson, Bond (1990), Requena-Silvente, Walker (2006),

Some product attributes can be additive combinable.[21] For example, price for each additional kilometer can be higher for business class than for the economy class.

3.3 Second step. Consumers' willingness to pay

According to the Rosen the second step is to incorporate the consumers and firm individual characteristics. We assume that all the firm face the same costs and the same production function, so there is no need to estimate an additional function for companies.

As far as consumers are concerned, the utility maximization problem provides us with an additional equation. In case of continuous choice version the consumers willingness to pay is:

$$\frac{\partial p(x)}{\partial x} = -\left(\frac{\partial u(x,z,p(x))}{\partial x} / \frac{\partial u(x,z,p(x))}{\partial p}\right) \tag{3}$$

where z_i is a vector of the individual characteristics for each consumer.

With the quasi-linear utility function $u(x,z,p(x)) = u(x,z) - p(x)$, the equation (3) is simplified to:[22]

$$\frac{\partial p(x)}{\partial x} = \frac{\partial u(x,z)}{\partial x} \tag{4}$$

In case of discrete choice version, when the product attributes are not continuous, and also for quasi-linear utility function consumers willingness to pay to x_k over x_n is

$$p_k - p_n = u(x_k, z) - u(x_n, z) \tag{5}$$

So knowing the functional form of the utility we are able to estimate the equation (4). For simplicity we assume linear form of the utility function: [23]

[21] Chan (2006)

[22] Quasi-linear utility function is used in a large number of the articles on the similar topics: Berry (1994), Berry and all (1995), Berry and all (2004), Nevo (2000), Ekeland and all (2004), Neigheim (2006)

[23] Linear form of the utility function is used by Berry (1994), Berry and all (1995), Berry and all (2004), Nevo (2000)

$$u_{i,j,t} = \sum_k x_{j,k}\widetilde{\beta}_{i,k} - p_{j,t} \tag{6}$$

The coefficient $\beta_{i,k}$ is consumer specific, which means each consumer has her own individual value for one or another product attribute. And this individual attribute valuation depends on clients personal characteristics:[24]

$$\widetilde{\beta}_{i,k} = \bar{\beta}_k + \sum_r z_{i,r}\beta^0_{k,r} + \nu_{i,k} \tag{7}$$

The dependent variable can be found from the equation (4):

$$\widetilde{\beta}_{i,k} = \frac{\partial p(x'_j)}{\partial x_k}, \tag{8}$$

where both sides represent consumers' willingness to pay for product characteristic k.

In case of discrete choice model willingness to pay also can be estimated for different types of discrete variables.

Connected through utility maximisation problem (equation 8) the equation 1 and equation 7 are estimated as a simultaneous system. So, finally we will be able to analyze the influence of the different product attribute on the prices and analyze the influence of the consumers' characteristics on their willingness to pay.

4 Data

For the purpose of this analysis we use the Survey of International Air Travelers conducted by the American International Trade Administration Office of Travel and Tourism Industries. This survey took place in several airports of the U.S. The passengers were asked series of questions about their trip. First of all, the respondents gave the details about the flight: airline company, destination point, departure date, etc. All the flights are international: from the U.S. to abroad.

[24]see Berry and all (2004)

If a passenger has any connecting flights, the survey provides information about the starting points and principal destinations of their travel. Secondly, there is also a part of the survey dedicated to the personal characteristics of the consumer: country of origin, age, income, etc. Moreover there is also a number of the additional questions: the purpose of the trip, with whom the respondent is traveling, hotel references, etc. Finally, the passengers were asked about their expenses that gives us the most interesting parameter: ticket price. They also gave their opinion about quality of all the types of the provided services.

This database is unique in a way that it provides ticket prices and dates of the departure. The respondents were also asked how many days prior to the departure the ticket had been purchased, which gives us a possibility to control for the reservation time. The combination of these three variables cannot be found in the databases used in the previous research.[25] The previous articles devoted to the airline competition research employ other databases for the estimations. The commonly used is Destination and Origin Survey (DB1B, by Department of Transportation), a 10% random sample of the all tickets sold in the U.S.[26] This database is highly reliable but it does not include any details about the passengers or product characteristics. It also concerns the other databases used for competition analysis of the airline market before: Digest Statistics (ICAO), World Air Transport Statistics (IATA), etc.[27]

Although the Survey of International Air Travelers is a large database, it does not contain all the information necessary for our analysis. For examples, the flight distance is an important attribute. We use geographical distance between two points, not the number of flown kilometers, in order to have the same distance for every couple origin-destination points. From the consumers' point of view they have to cover the distance from one point to another regardless of how many kilometers the aircraft flies. The Distance variable is taken from another American database - All Carrier Statistics, T-100 International Market - a highly reliable database of Bureau of Transportation Statistics (all the data are received directly from the airlines). The T-100

[25] for details see section 1 Introduction

[26] see Borenstein (1989), Borenstein and Rose (1994), Dresner and Tretheway (1992), Evans and Kessides (1993), Brueckner and Spiller (1994), etc.

[27] see Gagnepain and Marin (2006, 2010), Marin (1995), Oum and all (1996)

International Market database has a variable called "Distance", which provides the geographical distance between two airports.

The social-economic characteristics of the cities are thought to influence the ticket price, therefore several more variables are added: GDP per capita of the destination country and the population of the origin and destination cities.[28] There is no need to include GDP of the origin country as it is always the U.S. GDP data are taken from the World Bank database, population from the site www.citypopulation.de.

As it was mentioned above the survey we employ has not been used for competition analysis before. It is can be explained by a number of disadvantages of the database. First of all, it is survey type data so there is a great number of missing information. Secondly, supposedly, there are mistakes in variables. For example, not all the respondents remember the exact day of the ticket purchase, so they give an approximate number of days. The same problem can arise with the most important variable for our analysis, the ticket price. The respondent can make a mistake providing the information about the ticket price, so we don't possess the true value of the ticket price (P_t^{obs}); we have a price which can contain a possible measurement mistake ($P_t^{obs} = P_t^{real} + \varepsilon_t$). The third disadvantage of the Survey is a sample problem. We cannot be sure that this database is a perfect approximation of the real market. For example, it is not possible to obtain any plausible information about the volumes of production. Nevertheless, the similar databases are used for travel demand analysis.[29] We employ this database regardless all the disadvantages, as only survey type of data can provide us with all the ticket details and consumers' characteristics.

4.1 Data handling and descriptive statistics

A huge amount of work has been done in order to prepare the database for the necessary estimations. We start the analysis with the definition of a route. The principal characteristic

[28] see Berry (1990), Brueckner and Spiller (1994), Marin (1995), Peteraf (1995), Zhang and Park (1998), Ciliberto and Tamer (2009), Forbes and Lederman (2009), others

[29] see I. Introduction

of a route is the origin and destination points. The database provides codes of the origin and destination airports of an international direct flight originated from the U.S. But there are a lot of passengers traveling by indirect flights and we have to take them into consideration as well. During the survey respondents answered the questions about the starting and destination points of their journey. The database provides the origin airports which are situated on the territory of the United States. As far as destination point is concerned, we know only the city (town) or region of destination, not the airports. Therefore it was decided to use the cities or towns as the origin-destination points. The lack of the information does not allow us to use the airport in our analysis. Moreover, this approach allows having more observations for every route that could make our results more consistent.[30]

After the unification of the origin and destination information, we receive a database of a number of direct and indirect flights covering the whole journey of a passenger. The number of coupons vary from one (for direct flight) to three. In database we have only variable that indicates what airline served the direct international flight, so we have to assume that the same airline which is indicated for the direct flight served all the connecting flights. It is a reasonable assumption, as it is usually much cheaper to buy the whole trip from the same company than to buy the tickets separately from point to point.[31] Some of the flights were bought as a part of the package tour.[32]

Then we have to delete a number of observations for several reason. First of all, as in any survey there is a lot of omissions when the interviewees failed to answer some questions, so we eliminated the observations without information for crucial variables, like price or reservation moment. Two more types of observations were also deleted: one way tickets and first class

[30]The both approaches are used in the literature. Some authors use the airports not cities as the origin-destinations points (Borenstein (1989) Morrison, Winston (1990), Kim, Singal (1993), Borestein, Rose (1994)). Others prefer cities pairs for the analysis of competition on the airline market (Berry (1990), Forbes and Lederman (2009) Barley, Pazar (1981), Dresner, Tretheway (1992), Evans, Kessides (1993) Brueckner, Spiller (1994), Sinclair (1995), Marin (1995), Oum, Park, Zhang (1996)). There is also a number of article with the control for both: cities and airports (Brueckner, Dyer, Spiller (1992),Ciliberto and Tamer (2009)).

[31]Sometimes the tickets for an indirect flight are more expensive than direct flights by the same company and for the same route, however they cost more for the company. It also a yeild managment strategy allowing to distinguish consumers more sensitive for price.

[32]In the models below we test if the fact that ticket was bought as a part of the package tour influences the price. We have not found any significant influence

tickets. For both of them the airline companies employ special pricing strategy and these tickets present a very small percentage of the data. Calculation of the descriptive statistics shows a huge dispersion of the price per kilometer, so in order to minimize the measurement mistake we deleted the extreme values: 5% quantile for the highest values and 5% quantile for the smallest price values. After the data cleaning, less than a half of all the observations has left.

Tables 2-3 present descriptive statistics of the main variables. There are different types of routes included in the database. All of them originate from the U.S., but they have various destinations points starting from cities in Canada and Mexico to Japan and South Korea. That gives us a huge dispersion in distance from 1,000 kilometers to 16,000 kilometers. There is also a large variety in Origin and Destination population: from 10 thousand people in Shannon, Ireland to huge metropolises like Shanghai and Istanbul with more than 10 billion people population.

As another factor for the model we used GDP per capita in country of destination and it is also has a large dispersion. Prices are various, which can be explained by the diversity of routes and their characteristics. Moreover, two ticket classes - business and economy - have also a serious effect on price dispersion. The price per kilometer is also calculated but it is not used in regression.

The maximum period of the reservation is 250 days in the restricted base. In the original base this figure is much bigger (730 days, but less than 1% of the tickets reserved more than 300 days before departure). The reservations made more than eight months before departure are usually exceptional.

75% of the tickets correspond to the economy class and 25 % of them are business class tickets. One half of the consumers travel by direct flights, the other with one or two stops. In reality the part of direct flights is less (In the unrestricted database this ration is: 43% of direct flights). All the major routes have at least several direct flights and the restricted database capture the biggest routes so direct routes part is slightly higher.

Table 1: Descriptive statistics

		Mean	Std. Dev	Min	Max	Obs.
Dist	Distance, km	8 596	3676	728	17 684	11831
PopO	Population of origin city/town, thousands people	3 447	3 242	2383	8 175	11831
PopD	Population of destination city/town, thousands people	4 224	3 813	9673	14 200	11831
GDP	GDP of the destination country per capita, US dollars	16 832	14 023	294	62 143	11831
Price	Price per person, US dollars	1 306	1279	46	11 000	11831
Price/Km	Price per kilometer	0.16	0.14	0.013	0.88	11831
Res	The time before departure the reservation was made	45.5	50.0	0	365	11831

Table 2: Descriptive statistics

Ticket class		Freq.	Percent
1	Business	2 448	20.69
0	Economy	9 383	79.31

FF	Freq.	Percent
No	10 941	92.5
Yes	890	7.5

4.2 Reservation moment and the personal characteristics of the consumers

It is possible to offer different prices for the same flight only if the different types of consumers tend to reserve their tickets at different moments. In this subsection we try to show the connection between the personal characteristic of the consumers and reservation time.

During the survey the respondents answered a question how many days prior to the departure the ticket had been bought. Using this information we create a discrete variable from 1 to 9 corresponding to reservation time. Separation for nine groups are made according to the recommendations attached to the data base. The second and third columns of the Table 4 demonstrate how these groups are created.

The same table (Table 4) shows the connection between personal characteristics of consumers and reservation moment. The statistics presented in the Table 4 are calculated using the unrestricted database. The $Income$ variable is a discrete variable from 1 to 11, group 1 represents the lowest income, group 11 – highest income group. The Sex variable is equal to 1 for men and 0 for women, the Bus variable is equal 1 for business class tickets and 0 for economy class tickets.

Table 4 shows that women and older people prefer to make their reservation as early as possible. In general men buy tickets more often than women. The majority of the business class tickets are bought from 3 days to two weeks before departure. Surprisingly there is a considerable share of business class among tickets bought more than half of a year before departure. It may be connected with some important business events or conferences. As far as income is concerned, it does not have a monotonic influence on the booking time. It can be explained by the fact that richer people have probably more business trip, reserving their tickets closer to departure but at the same time they travel more for leisure purpose and buy their tickets more in advance. The largest part of the respondents buys their tickets from two week to a month in advance. More than 75% buy their tickets in the last two months and only 3% of the consumers obtain their tickets more than half a year before departure.

Table 3: Descriptive statistics

Rest	ResTime (days)	Obs.	%	Variable	Sex	Income	Age	Bus
1	1-3 days	783	7.18	Mean	0.31	5.95	42.43	0.20
				Std.dev		3.06	11.49	
2	4-7 days	1100	10.14	Mean	0.27	6.49	42.99	0.27
				Std.dev		3.06	11.70	
3	8-14 days	1438	13.31	Mean	0.29	6.48	43.66	0.29
				Std.dev		3.14	12.07	
4	15-30 days	3181	29.93	Mean	0.36	6.18	43.62	0.22
				Std.dev		3.10	13.24	
5	31-60 days	2051	19.22	Mean	0.43	6.02	44.32	0.15
				Std.dev		3.05	13.93	
6	61-90 days	1105	9.39	Mean	0.47	5.99	44.26	0.13
				Std.dev		3.01	14.79	
7	91-120 days	476	4.63	Mean	0.44	6.16	45.05	0.16
				Std.dev		3.03	15.41	
8	121-180 days	494	4.58	Mean	0.49	6.38	48.31	0.16
				Std.dev		3.11	14.69	
9	180 or more	173	1.62	Mean	0.63	6.91	48.54	0.21
				Std.dev		3.11	14.56	

5 Estimation results

This section is devoted to the estimation results. For the purpose of our analysis, a system of the nonlinear equations is estimated simultaneously.

$$lnPr_t = \beta_0 + \beta_1 * lnResT_t + \beta_2 * Bus_t + \beta' X_t + \alpha_t + \delta_t + \epsilon_t \quad (9)$$

$$WTP_t = \gamma_0 + \gamma' Y_t + \lambda' Z_t + \vartheta_t \quad (10)$$

where $lnPr_t$ - logarithm of the price for the ticket, $lnResT_t$- logarithm of number of days prior to the departure the ticket bought, Bus_t - dummy variable for business class, X_t- vector of control variables (route and ticket characteristics), α_t- company fixed effects, δ_t- fixed effects for some routes (biggest),[33] Y_t - vector of consumer personal characteristics, Z_t- vector of the ticket characteristics (controls), ϵ_t,ϑ_t- error term. Vector of control variables X_t: geographical

[33] The number of route is too large to include all of them in the estimation

distance between origin and destination points (Dist), population of the origin and destination points (PopO, PopD), dummy variable if ticket has been bought with the help of the Frequent Flyer program (=1 if Yes), GDP of the destination country, month and weekday fixed effects, holiday dummies. Vector of consumer personal characteristics, Y_t : age of the consumer (Age), income group (Income), number of international trips made for the last 12 months (Trip12m), fixed effects of the main purpose of the trip, fixed effects of the consumer's principal occupation. Vector of ticket characteristics (controls), Z_t - $lnDist$, $lnPopO$,$lnPopD$, $lnResT$

We also tried another specification of the first function in the system:

$$lnPr_t = \beta_0 + \beta_1 * lnResT_t + \beta_2 * Bus_t + \beta' X_t + \gamma' X_t * Bus_t + \alpha_t + \delta_t + \epsilon_t \qquad (11)$$

Estimating this function we try to check the theory that influence of the same factors is different for business and economy class. For example every additional kilometer can have larger influence on price for business class than for the economy.

As it was mentioned before there was no theoretical specification for a hedonic price function. So before proceed with the estimation of the system (equation 9; equation10) we run several tests in order to chose the best specification of the model. For these tests the equations are estimated separately. The previous chapter shows us that data is very heterogeneous, so it is necessary to control for the route, company, season influence. Therefore first of all we need to treat these effects. The standard approach to panel data estimation (our data can be treated as a panel of routes or companies) implies testing model without effects against model with fixed and random effects. Made for route and companies effects F tests and Hausmann tests mostly support the model with fixed effects.[34] Moreover, the employment of the standard fixed effects is commonly used in all types of analysis. Therefore we introduce route, company and some seasonal fixed effects in the nonlinear system. Secondly, there are still some doubts about the linear specification. So we try to test the linear and loglinear specification against specifications with higher power members using the RESET test. It shows that introduction of the quadratic

[34] There are a few doubtful results of Hausmann test for some spesifications.

or higher power members does not improve the specification.

After specification tests we proceed with the system estimation. The results of estimation of a system of nonlinear equations are presented in Tables 5-6. For this purpose we employ a standard method: two-step FGNLS estimator. There are four models presented in both tables with a slightly different specifications. Model 1 - Model 3 have different fixed effects. In Model 4 is devoted to more detailed analysis of business class influence, namely estimation of the equation 11. It is necessary to keep in mind that Tables 5-6 present two parts of the same model. The results are stable in all the chosen specifications which indicates good quality of the model. TheR^2 in the price regression is about 50%, so the chosen ticket characteristics are able to explain half of the price dispersion. The signs of the coefficients are expected. The increase of the reservation time by 1% leads to price decrease for about 0.04%. This figure means that when ticket is reserved three months in advance one days does not play a role, but if ticket reserved three days before departure every day can slightly change the price. However the received coefficient is smaller then we expected. Business class tickets are 80% more expensive than the similar tickets of economy class. Increase in route distance by 1% means increase in price by 0.6% in average. In average ticket bought with the advantages of the frequent flyer program have a 50% discount. GDP of the destination country and population of the origin and destination points has no significant influence on price. We expected to obtain the positive coefficient for population and GDP, as larger and economically developed cities produce much more demand for air service. But the obtained results have also a logical explanation: these kind of cities also provide a huge supply. Moreover for the airline companies working on a huge route can be more efficient than a small route. There are also touristic destinations like small islands, which can be very expensive regardless population and GDP.

The results of the Model 4 show us that indeed for some factors the difference between business and economy class is strongly significant. Every addition percent to distance increase price by 0.8% for business class comparing to 0.6% increase for economy class. Similarly, the

Table 4: Hedonic price function

Price	M1	M2	M3	M4
lnResT	-0,03***	-0,03***	-0,04***	-0,05***
	(0.005)	(0.005)	(0.005)	(0.005)
Bus	0.78***	0.80***	0.79***	-0.90***
	(0.01)	(0.01)	(0.01)	(0.31)
Bus*lnDist				0.19***
				(0.03)
Bus*FF				-0.15***
				(0.04)
FF	-0.48 ***	-0.49 ***	-0.48***	-0.40***
	(0.02)	(0.02)	(0.02)	(0.03)
lnDist	0.55***	0.59***	0.60***	0.58***
	(0.01)	(0.02)	(0.02)	(0.02)
lnPopO	-0.004	-0.003	-0.003	-0.002
	(0.004)	(0.004)	(0.004)	(0.004)
lnPopD	0.008	0.001	-0.002	-0.001
	(0.005)	(0.006)	(0.006)	(0.006)
lnGDP	0.012	0.009	0.006	0.006
	(0.004)	(0.006)	(0.006)	(0.006)
Xmas	0.15**	0.16**	0.10**	0.10**
	(0.06)	(0.06)	(0.06)	(0.06)
Easter	0.02	0.02	0.02	0.004
	(0.06)	(0.06)	(0.06)	(0.06)
Weekdays fixed effects			+	+
Months fixed effects			+	+
Big routes effects		+	+	+
Airline fixed effects	+	+	+	+
Constant	1.64***	1.43***	1.39***	1.58***
	(0.12)	(0.18)	(0.18)	(0.18)
Obs.number	9147	9147	9147	9147
R2	0.48	0.51	0.53	0.53

The standard deviations are in the parenthesizes. *- significant at 10% level. ** - significant at 5% level. *** - significant at 1% level.

Table 5: Reservation moment, WTP

Price	M1	M2	M3	M4
Sex	0.51***	0.66***	0.80***	0.82***
	(0.20)	(0.24)	(0.29)	(0.29)
Income	0.19***	0.22***	0.28***	0.29***
	(0.04)	(0.05)	(0.06)	(0.06)
Age	0.009	0.014	0.017	0.018
	(0.007)	(0.010)	(0.012)	(0.012)
Trip12m	0.04*	0.07*	0.08*	0.08*
	(0.03)	(0.04)	(0.04)	(0.04)
Controls	+	+	+	+
Occupation effects		+	+	+
Purpmain effects		+	+	+
Constant	-1.46	-4.34*	5.20*	-5.17*
	(1.95)	(2.39)	(2.90)	(2.93)
Obs.number	9147	9147	9147	9147
R2 (within)	0.32	0.32	0.32	0.32

The standard deviations are in the parenthesizes. *- significant at 10% level.
** - significant at 5% level. *** - significant at 1% level.

discount of the frequent flyer program is higher for business class by 15 percentage points. This fact has several logical explanations. First of all, a lot of business class consumers travel for business purpose, so they usually make trips more often, especially long distance international trips, therefore they are more suitable for frequent flyer program. Secondly, one of the popular approaches of the yield management is to upgrade the frequent flyers from economy to business class. On average in Model 4 the business class is more expensive for about 80% (on average means for distance of 8 500 and full price payed). More detailed analysis of the business class influence does not change the results dramatically and has almost no impact on the quality of the regression.

Finally, the coefficients for fixed effects are not presented in the table but the estimated values are also expected. The tickets are most expensive during the summer months and in December: about 20% more expensive during these four months. For other months the difference between them is not statistically significant. The cheapest tickets are for Tuesday and Wednesday, the most expensive ones are for Saturday.

Table 6: Busness class, WTP

Price	M1	M2
Sex	-17.9***	-18.4***
	(6.93)	(6.92)
Income	-3.35***	-3.05**
	(1.15)	(1.15)
Age	-1.36***	-1.33***
	(0.29)	(0.29)
Trip12m	-3.78	-11.67***
	(1.12)	(2.49)
Trip12m^2		0.47***
		(0.13)
Not 1st trip	10.7	4.86
	(15.0)	(15.1)
Controls	+	+
Occupation effects	+	+
Purpmain effects	+	+
Constant	-5900***	-5900***
	(102.6)	(102.5)
Obs.number	9147	9147
R2 (within)	0.67	0.68
F statistics	373.1	359.6

The standard deviations are in the parenthesizes. *- significant at 10% level. ** - significant at 5% level. *** - significant at 1% level.

It is worth to mention that we tried to introduced more of the control variables in the regression (5) but the coefficients of these variable are not significantly different from zero. [35]

The second regression of the system is devoted to the analysis of connection between consumers' willingness to pay for thier flexibility and certain personal characteristics. Dependent variable is consumers willingness to pay for one additional day of their flexibility, which means how much more they are ready to pay for a ticket to have the right to buy it one days closer to the departure date. In average this variable is equal to 4.5 dollars per days. The results presented in the Table 6 show that the quality of the regression is quite high, R^2 is about 30%. The gender of the consumers plays a significant role. In average the women are ready to pay more for their flexibility: 0.6 - 0.8 dollar plus for every additional day. These results are surprising because statistically women buy tickets earlier than men. Rich people also appreciate their flexibility: every next income group tends to pay 0.3 dollar more for every additional day. Age has no significant influence. We also tested quadratic influences of age but we failed to find any significant impact. One additional trip per year increase the willingness to pay for flexibility by 0.08 dollar. Occupation and purpose effects have expected significant influence on willingness to pay.

The last equation of our model is willingness to pay for business class. For average ticket of 1000 dollars business class customers pay 800 dollars more. Women's willingness to pay for business class is less than men's, but the difference is not huge: about 18 dollars per ticket. Surprisingly consumer income has negative significant effect on willingness to pay for business class. Age has also a negative impact: elderly people appreciate business class less: the willingness to pay decrease by 1.5 dollar with every year of the consumer. Influence of number of trip is negative but quadratic. Occupation and purpose effects are jointly significant and thye have expected signs.

The developed model covers two main instruments of yield management: moment of purchase and ticket class. Unfortunately we do not have a possibility to control for the last principal

[35] Number of coupons, length of trip, number of visited cities

instrument of yield management: willingness to pay for frequent flyer program bonuses.

6 Summary, Conclusion and Proposed extensions

This article presents an extensive analysis of the pricing policies on the airline market. Our research is devoted to the analysis of ticket price heterogeneity. The airline market is shown as a market with a differentiated product. We manage to estimate a price function taking into consideration both sides of the market: firms' strategies and consumers' personal preferences. A special attention is paid to the moment of reservation as it is the main instrument of the airline pricing policies. This analysis becomes possible only with a new database: Survey of International Air Travelers which allows us to obtain all the necessary variables for the estimation.

The first part of a system of theoretical equations is a hedonic price function. It explains the connection between product characteristics and price. The obtained results are expected. The earlier reserved tickets cost less. The business class is more expensive in average for 80%. The tickets are also more expensive during the summer months and just before Christmas. More detailed analysis of the business class influence does not change the results dramatically and has almost no impact on the quality of the regression. However it has introduced a couple of new significant coefficients.

The second part of the estimated system is devoted to the consumers willingness to pay for certain ticket attributes. The estimations show that there is a significant connection between personal characteristics of the consumers and their buying patterns. Firstly, consumers' willingness to pay for their flexibility is estimated. The women buy their tickets more in advance but at the same time they are ready to pay more for their flexibility. The richer people also prefer to stay more flexible. If consumers travel a lot they also tend to pay more for a possibility to buy ticket later. We found no significant influence of the age on the preference of the reservation moment.

The willingness to pay for business class tickets is also investigated. Women appreciate business class less than man. The more passengers travel less they are ready to pay for the business class, however the influence is quadratic. Surprisingly the model finds negative significant connection between income and consumers willingness to pay for business class.

The conducted analysis can be extended in several different ways. First of all, there are still some specification, database, unobserved variables problems. More data can allow us to solve some of the problems. Secondly, the willingness to pay can be measured separately for different tickets types, which can improve the quality of the estimation. Thirdly, within the framework of this paper we plan to develop an explicit recommendation to the airline companies based on the obtained results. Having more concrete information about consumer preferences airline companies can develop more flexible system of discounts for different groups of customers.

References

[1] Borenstein, Severin "Hubs and High Fares: Dominance and Market Power in the U.S.", *The RAND Journal of Economics*, vol 20, No 3 (Autumn 1989)

[2] Borenstein, Severin, Nancy L. Rose "Competition and Price Dispersion in the U.S. Airline Industry", *Journal of Political Economy*, 1994, vol. 102, No.4, pp.653 -683

[3] Botimer, T.C. and P. P. Belobaba "Airline Pricing and Fare Product Differentiation: A New Theoretical Framework", *The Journal of the Operational Research Society*, Vol. 50, No. 11, (Nov., 1999), pp.1085-1097

[4] Brueckner Jan K., Pablo T. Spiller "Economies of Traffic Density in the Deregulated Airline Industry", *Journal of Law and Economics*, Vol.37, No.2 (Oct, 1994), pp.379-415

[5] Chen, Victoria C. P., Dirk Günther and Ellis L. Johnson "Solving for an Optimal Airline Yield Management Policy via Statistical Learning", *Journal of the Royal Statistical Society*, Series C (Applied Statistics), Vol. 52, No. 1(2003), pp.19-30

[6] Clemons, Eric.K, Il-Horn Hann and Lorin M. Hitt "Price Dispersion and Differentiation in Online Travel: An Empirical Investigation", *Management Science,* Vol. 48, No. 4, April 2002 pp.534–549

[7] Coughlan, J "Airline Overbooking in the Multi-Class Case", *The Journal of the Operational Research Society*, Vol. 50, No. 11, (Nov., 1999), pp.1098-1103

[8] Dana, James D.Jr "Using Yield Management to Shift Demand When the Peak Time is Unknown", *The RAND Journal of Economics,* Vol. 30, No. 3 (Autumn, 1999), pp.456-474

[9] Desiraju, Ramarao and Steven M. Shugan "Strategic Service Pricing and Yield Management", *Journal of Marketing*, Vol. 63, No. 1 (Jan., 1999), pp.44-56

[10] Dresner Martin, Michel W. Tretheway "Modeling and Testing the effects of Market Structure on Price", *Journal of Transport Economics and Policy*, Vol. XXVI, No2, May 1992, pp.171-184

[11] Evans William N., Kessides Ionnis N. "Localised Market Power in the U.S. Airline Industry", *The Review of Economics And Statistics*, vol.75, No.1m (feb. 1993), pp.66-75

[12] Feng, Youyi and Baichun Xiao "A Continuous-Time Yield Management Model with Multiple Prices and Reversible Price Changes", *Management Science*, Vol. 46, No. 5 (May, 2000), pp.644-657

[13] Gagnepain, Philippe, and Pedro L.Marin "Regulation and Incentives in European Aviation", *Journal of Law and Economics*, vol. XLIX (2006)

[14] Gagnepain, Philippe, and Pedro L.Marin "The Effects of Airline Alliances:What Do the Aggregate Data Say?" , SERIEs, Spanish Economic Association, vol.1 (3), July 2010, pp.251-276,

[15] Giaume, Stephanie and Sarah Guillou "Price Discrimination and Concentration in European Airline Markets", *Journal of Air Transport Management*, Vol. 10 (2004), pp.305–310

[16] Leal de Matos, P. "Yield Management for Privatised Air Traffic Control?", *The Journal of the Operational Research Society*, Vol. 52, No. 8, (Aug., 2001), pp.888-895

[17] Smith, Barry C., John F. Leimkuhler and Ross M. Darrow "Yield Management at American Airlines", *Interfaces*, Vol. 22, No. 1, (Jan. - Feb., 1992), pp.8-31

[18] Leal de Matos, P. "Yield Management for Privatised Air Traffic Control?", *The Journal of the Operational Research Society*, Vol. 52, No. 8, (Aug., 2001), pp.888-895

[19] Marin Pedro L. "Competition in European Aviation: Pricing Policy and Market Structure", *The Journal of Industrial Economics*, Vol. 43, No.2 (june 1995), pp.141-159

[20] Peteraf Margaret A. "Sunk Costs, Contestability and Airline Monopoly Power", *Review of Industrial Organisation*, Vol.10, 1995, pp.289-306

[21] Stavins (2001) "Price Discrimination in the Airline Market: The Effect of Market Concentration"*, The Review of Economics And Statistics*, vol.83, No.1, (feb. 2001)

[22] Zhang Anming, Park Jong-hun "An Empirical Analysis of Global Airline Alliances: Cases in North Atlantic Market" August 1998

[23] Zhang Anming, Park Jong-hun, Oum Tae Hoon "The Effects of Airline Codesharing Agreements on Firm Conduct and International Air Fares", *Journal of Transport Economics and Policy*, vol.30, no. 2. (may, 1996), pp.187-202