Department of Economics and Finance

Dynamic Pricing in the Airline Industry

Bachelor Thesis

Author: *Supervisor:*

Giorgenzo Treves de Bonfili . Prof. Marco Dall'Aglio

ID Number: 198461 *Subject:* Mathematics 1

A.Y. 2017-2018

June 11th 2018

ABSTRACT

An often neglected, albeit fundamental, aspect of the airline pricing issue consists in determining, from the customers 'point of view, how companies assign a fare to all the available seats on an airplane in order to maximise their revenues. This work aims at presenting how flights fares are determined in practice, and how their distribution changes over time. Focusing on Easyjet, a low-cost carrier, the work aims at determining the algorithm associated with flight fares computation and the variables the system considers in order to maximise the revenues the company earns. To understand how this algorithm works, Easyjet use case is considered and its revenue maximisation strategy under Bellman's optimality discussed. Under some assumptions, such as uniform purchase probability and monopolistic competition, a dynamic pricing algorithm as well as a pricing model are presented and examples of the necessary data collection in a real scenario are drawn. By blending theory and practice, I question how prices are determined in practice and which factors have a greater role in affecting the seats' price distribution. Indeed, with respect to fare statistical distribution, three main factors appear to affect how prices evolve over time. Firstly, fare distributions are increasing over time because the company enacts price discrimination to fully capture the willingness to pay of consumers with different needs. In particular, people travelling for business tend to buy their tickets few days before departure and their price elasticity function is approximately inelastic. Secondly, over time fare distributions move, on average, downward to reflect the perishable nature of a flight's seat. Thirdly, to avoid strategic behaviour from consumers who tend to postpone the purchase of their tickets, the price observed by prospective buyers tends to increase as the date of departure nears. In the light of the work's claims, a model of the algorithm used by major low cost airline companies is developed and validated while further applications in other industries are suggested.

SUMMARY

1. INTRODUCTION

Algorithm-based reservation systems were developed in the 1950s to keep track of airline seat booking and fare information. Initially, these were internal systems but they were soon made available to various agents in the market. The deregulation of airline pricing in the following years allowed a much broader use of these systems for economic activity, especially related to pricing. Indeed, the complexity of airline pricing has grown over the last decade. Nowadays, the revenue management systems for pricing (RM) employed by airline companies, such as Easyjet, Ryanair and Wizzair, have turned it into one of the most customized and hidden yet fascinating processes in the market. Indeed, this pricing practice is obscure vis-a-vis both customers and competitors, that is, neither the customers nor the competitors know how the dynamic pricing algorithm is actually implemented in the market. On one hand, the first could be tempted to switch to another company because of the price discrimination they have to bear while, on the other hand, the latter could adopt strategies of adversarial marketing and, thus, possibly decrease the company's revenues. Currently, airline dynamic pricing represents a great challenge for modern economic studies because it is so distant from classic models and assumptions as well as from their level of analysis.

Dynamic pricing (DP) in airline markets is related to the way fares on sale evolve over time (McAfee & te Velde, *Dynamic Pricing in the Airline Industry*, 2007). According to McAfee's theoretical model, three major factors are assumed to shape the fares' temporal profile and, thus, contribute to its complexity. First, airlines sell a highly perishable service. This means that while the fare today has to account for the cost of the foregone option of selling the seat later on for a higher fare, for flights such option value goes to zero as the take-off approaches. Therefore, it is reasonable to predict fares falling over time (McAfee & te Velde, *Dynamic Pricing in the Airline Industry*, 2007). On the other hand, carriers, to maximise profits, may want to discriminate the business passengers' segment from other demand travellers, e.g., tourists. As the former are more likely to learn about their need to travel only a few days before the departure date and their demand is quite inflexible, fares are expected to rise over time (Gaggero and Piga, *Airline Market Power and Intertemporal Market Dispersion*, 2011; Alderighi, Nicolini, & Piga, *Targeting leisure and business passengers with unsegmented pricing*, 2016). Third, a similar increasing fares profile can emerge when customers are strategic and may postpone the purchase seeking last-minute discounts (Deneckere and Peck, *Dynamic Competition with Random Demand and Costless Search*, 2012; Sweeting, *Dynamic Behaviour in Perishable Goods Market*, 2012). A commitment to raise fares over time is often necessary to discourage such attitude, unless the probability of a stock-out is high (Moller and Watanabe, *Advance Purchase Discount and Clearance Sales*, 2010).

To explain the dominant trend of the fare profile (i.e. whether the price is an increasing or decreasing function of time), in this work I will abandon the analysis based on a single fare so far broadly used in the literature and adopt the notion of fare distribution developed by Dana *(Equilibrium Price Dispersion under demand uncertainty: the roles of costly capacity and market structure,* 1999). According to Dana, the airline pricing algorithm does not only define the fare of the seat on sale, but also of all the remaining seats on the flight. The research question appears to be: "How are fares assigned and how does their distribution change over time?" The innovative aim of this work is to guide the average consumer in the decision-making process and minimise the cost she has to pay. In order to tackle this issue I have taken the carrier's perspective and proceeded through some important steps. First, I have shown the evolution of the literature on dynamic pricing that describes how such distributions are shaped. In particular, the current view is that airline companies arrange seats into groups, denoted as "buckets", where each bucket is defined by an increasing price tag and a variable size. Second, by suitably referring to a model designed by Alderighi, Gaggero, & Piga, (*The hidden side of dynamic pricing in airline markets*, 2016), I will present a theoretical model of dynamic pricing. In particular, I will show that each seat in the distribution is affected by both an intrinsic declining value and an extrinsic increasing value due to several factors. While the total effect could appear ambiguous, empirical evidence collected in Sector 4.4 will show that the latter factors seem to prevail as the departure date approaches. Moreover, albeit strong, the monopolistic assumption will be explained thoroughly in the paragraph concerning airline revenue management. Third, through the development of a dynamic pricing algorithm and the characterization of distributions at a flight's level obtained by a careful observation on Easyjet's web-based reservation system, it is possible to determine how DP is implemented in practice. In brief, I will show graphically that DP associated with changes in the bucket sizes is quantitatively more relevant than changes involving modification of the buckets' fare levels which conversely tend to remain rather invariant over time. In other words, once the pricing function has been discretized, the different levels of discretization adopted will not vary in time.

The remainder of the paper is structured as follows. In Section 2 a literature review on dynamic pricing is presented and linked to the different views on the topic. Moreover, a disclosure on airline revenue management is given and a problem statement made. In Section 3, under some assumptions and caveats, a theoretical pricing model is presented, drawing on both previous studies as well as on real world scenarios. The development of a dynamic pricing algorithm is explained in Section 4. The different attempts to build this algorithm are discussed and a proof of concept on a major Easyjet route, Rome-Amsterdam, is provided. In section 5, the results of the experiment conducted in Section 4.4 are explained and an analysis on the possible usage of my algorithm as a proxy of Easyjet's real one is made. In other words, I will question how good is my algorithm in predicting the fare distribution of the route selected and how its efficiency can be improved.

Furthermore, by referring to actual data, I will graphically show the bucket intertemporal evolution for two Easyjet flights with different departure dates on the same route. In conclusion, a final assessment of the situation in the transportation industry is made and limitations of the model as well as further possible applications are suggested.

2. LITERATURE REVIEW ON DYNAMIC PRICING

2.1 Previous studies on dynamic pricing

Dynamic programming appears to be very useful in the airline industry. Indeed, airfares are determined, among other factors, by dynamic adjustment to stochastic demand given limited capacity. Airlines also adjust prices on a day-by-day basis as capacity is limited and the future demand for any given flight is uncertain. While fares generally increase as the departure date approaches, prices can actually fall from one day to the next, after a sequence of low demand realizations. This pricing strategy is known as dynamic pricing.

Dynamic pricing (DP) is a rather broad concept encompassing several approaches in the academic literature. Although it usually encompasses any change in prices occurring over time, its diverse definitions appear to be a consequence of the different theoretical and empirical approaches developed to consider the pricing behaviour of firms. DP is often associated to a price change that is directly linked to at least one factor or event that induces a revision of the followed pricing procedure. For instance, in the sport industry, the decreasing prices of Major League Baseball tickets in secondary markets in Sweeting (*Dynamic Behaviour in Perishable Goods Market*, 2012) represent a famous indication of an active DP intervention by sellers in the form of the decision to relist the tickets at a lower price. Conversely, in Abrate et al. (*Dynamic Pricing Strategies: evidences from European hotels*, 2012) hotel rooms prices are found to be either increasing or decreasing over the booking period for stays during, respectively, weekends and weekdays. While these different findings certainly denote distinct "inter-temporal pricing" profiles, they cannot be unambiguously classified as instances of DP in terms of the definition mentioned above since they result from an empirical model where the source of price variation over time is not specified. For example, hotels may have determined them at the start of the booking season, and the decreasing or increasing profiles may be the result of a purely time-invariant pricing approach.

In the airline market, pricing policies are central for any empirical analysis. Borenstein and Rose (*Competition and Price Dispersione in the US Airline Industry,* 1994) distinguish between systematic and stochastic peak-load pricing as sources of price dispersion in the market. In the former, the price variation is based on systematic, that is, foreseeable and anticipated, changes in shadow costs known before a flight is available for booking, while the latter reflects a change in the probability during the selling season that demand for a flight exceeds its capacity. Most conspicuously, the distinction in Borenstein and Rose (*Competition and Price Dispersione in the US Airline Industry*, 1994) can be related to carriers' specific revenue management (RM) activity, intended as a process of both setting ticket classes as well as defining the number of seats available at each fare. Thus, RM encompasses both a systematic and a dynamic pricing dimension: the former can be seen as the outcome of the process before a flight enters its booking period, the latter represents subsequent changes over time to the initial composition of ticket classes, both in terms of fare levels and number of seats available in each class.

By capturing this two joint approaches, Dana (*Equilibrium Price Dispersion under demand uncertainty: the roles of costly capacity and market structure*, 1999) illustrates how, in a theoretical model with demand uncertainty and high capacity costs in case of an empty plane , it is optimal for airline companies to commit to an increasing fare distribution. In this case, each fare reflects the fact that the shadow cost of capacity is inversely related to a seat's probability to be sold. According to Dana's model, the fare charged should reflect the ranked position of the seat on sale in a step chart fare distribution. To implement such a graph, it is necessary to know the plane capacity at the time *t* a fare is either posted online or a ticket is sold. All the works on this issue provide evidence in support to the hypothesis of fares increasing as a flight fills up. Interestingly, Alderighi et al. (*Targeting leisure and business passengers with unsegmented pricing*, 2015) derive the same results by using two fares, the seat on sale and the last seat in the distribution; their approach is further developed in the present work, where the fare distribution of all the seats will be modelled as a step chart.

Since in Dana *(Equilibrium Price Dispersion under demand uncertainty: the roles of costly capacity and market structure*, 1999) firms cannot change the initial distribution they set, the model cannot provide any theoretical prediction on how firms would modify the price distribution over time. Would all fares start low and then increase or start high and then decrease? The question of the optimal temporal profile of fares is thus generally addressed in the operational research literature surveyed in Talluri and van Ryzin (*The Theory and Practice of Revenue Management*, 2004) and in McAfee and te Velde (*Dynamic Pricing in the Airline Industry*, 2007). In both fare setting models the focus is on the opportunity cost of selling one unit of capacity, i.e., the value "not-to-sell" the unit today and reserve it for a future sale. As shown also in Sweeting (*Dynamic Behaviour in Perishable Goods Market*, 2012), the value of the option "not-to-sell" is expected to fall over time, leading to a similar prediction for fares. In the theoretical model of Section 3, referring in particular to the findings of Alderighi and Piga, (*The hidden side of dynamic pricing in airline markets*, 2016), I show that if airlines can revise the fare distribution more than once, then under standard assumptions of demand, customers' evaluations and arrival rates being constant over time, the fares of all the seats would be expected to decline over time. Nevertheless, as already mentioned in the introduction, there are at least two reasons proposed in the airline literature as to why consumers face increasing fares over time. First, offering advance-purchase discounts can be an optimal strategy when both individual and/or aggregate demand is uncertain (i.e. individuals learn their need to travel at different points in time and airlines cannot predict which flight will enjoy peak demand), and consumers have heterogeneous valuations. Second, the revenue management models that predict a declining option value assume a constant distribution of willingness to pay, and therefore do not account for the fact that business travellers tend to book at a later stage (Alderighi et al., *The hidden side of dynamic pricing in airline markets*, 2016).

The present work does not aim at distinguishing among competing theories in the airline markets. On the contrary, it wants to blend the theory on airline dynamic pricing strategies with a data analysis process, in order to provide customers with a tool to understand the price low cost carriers allocate to a given seat. Nevertheless, the external comprehension of how the pricing algorithm works is hindered by, among others, complex revenue management practices and information asymmetry on what dynamic pricing algorithm implementation is actually used. As Google complex page rank algorithm implementation is unknown even to professionals in the field, the above mentioned factors impede a complete knowledge to external users to the firm. To solve this issue, on one hand, I will build a data analytics pipeline through which I will predict, under certain assumptions, Easyjet's pricing trend. On the other, I will test my pricing algorithm model and measure the degree of precision to which "the walk matches the talk", that is, the effectiveness of my prediction.

2.2 Disclosure on the real world scenario: the theory and practice of Revenue Management (RM)

Revenue management is the application of analytical tools that predict consumer behaviour and optimize product availability as well as price to maximize revenue growth. In an effort to sell their goods at a price that is as high as possible, both firms and individuals have always resorted to price adjustments until a stable equilibrium is reached. In particular, the last decade has witnessed an increased application of scientific methods and software systems for dynamic pricing, especially in the optimization of pricing decisions.

While retailers use price-based RM, firms in the airline industry use quantity-based RM, with an important exception: low cost carriers. The difference between these two approaches boils down to the extent to which a company is able to vary quantity or price in response to changes in market conditions. In particular, EasyJet, the company I consider in the work, operates price-based RM as it perceived as preferable by the firm. Quantity-based RM, which works by rationing the quantity sold to different products, results in reducing sales by limiting supply. Conversely, if the carrier has price flexibility, it is able to reduce the quantity sold by increasing the price, rather than reducing sales by limiting supply. This achieves the same quantity-reducing function, but it leads to higher profits since, by increasing the price, EasyJet reduces sales *and* increases revenues at the same time.

Although it may sound trivial, at this point a caveat is needed: not all dynamic pricing involves simple price reductions. As already mentioned, EasyJet employs price-based RM but empirically we can see that prices go up over time. Why? The company offers a non-refundable, one-way fare ticket without advance-purchase restrictions. However, during the booking period for a generic flight, prices vary based on capacity and demand for that specific departure. Our experience teaches us that the earlier you book the cheaper the fare *should* be. Still, due to different markets forces, sometimes fares can apparently remain constant over time. An emblematic force is indeed popularity, that is how many people are expected to buy a ticket for a specific flight. This variability depends on both exogenous and endogenous factors. In fact, any dynamic-pricing model requires an idea of how demand, either individual or aggregate, responds to changes in price. An additional factor to consider concerns supply that depends on the state of the market condition, specifically the level of competition.

Speaking about the supply, an arguably realistic assumption adopted when building RM algorithms is that low cost companies are operating in a monopolistic market. A telling example that explains the rationale behind this apparently counterintuitive assumption is the following: if Easyjet decides to lower its fares, competitors, assuming no service differentiation occurs among them, respond by lowering their prices too. With lower prices, the firm and its competitors see an increase in demand. The increase in demand is treated empirically as a monopolistic demand response function to Easyjet's price change, although competition is indeed at work in the industry. Moreover, an algorithm that considers the company as operating in an oligopolistic market would be much more complex to deal with, as it would entail also the concept of strategy. Therefore, for the sake of simplicity, in the present work I will consider Easyjet as operating in a monopolistic market.

With respect to the demand, EasyJet's pricing algorithm assumes myopic customers (i.e. they buy as soon as the offered price is lower than their willingness to pay). One can argue that a strategic-customer model is more realistic. Nevertheless, this kind of model would set the pricing process as a strategic game between the customers and the firm and will result in a very difficult problem to deal with. Luckily, in many situations consumers are very spontaneous when making decision and strategic behaviour can safely be ignored. Furthermore, in contrast with the classical economic theory, they often do not have sufficient time or information to behave strategically.

2.2 Problem Statement: how to establish how prices change over time and why are reservation prices hidden?

It is difficult to understand airline fares distributions both because of internal revenue management practices as well as because of customers' information asymmetry with respect to which pricing algorithm is actually used by the company. For this reason, I claim that a possible solution could entail an analytical approach based on data science. In particular, in order to disentangle the hindrances of the unknown pricing algorithm used by Easyjet, I will harness a pipeline of analysis that starts from data retrieval on a number of Easyjet flights, and predicts the trend seat prices will follow over time.

In economics, the concept of reservation price refers to the minimum price the company is willing to earn from the sale of a given seat. Airline companies hide their reservation price, for among others, three main reasons.

First, airlines companies aim at maximizing revenue per flight. This means charging more or less over time depending on what they predict will maximize total revenue for the flight by using the already explained "revenue management" or "yield management". As it is now clear, it is not very profitable to leave many seats empty as the company may have charged too much and could have made more money by selling at a lower price. On the other hand, sometimes it is also not profitable for the plane to leave completely full as the company may have charged too little and is losing money on the trip. Consistent with the economic theory, by adopting a dynamic pricing strategy, the company will maximise its revenues.

A second reason is that the product (a seat on an airplane) is worth substantially more to some people than to others. There are all kinds of "travellers" to accommodate: last-minute business travellers who will pay a very high price for a convenient flight, book-way-in-advance vacationers who have gotten some time off and want to arrange a comfortable trip, leisure travellers who will take whatever is cheapest. The airlines have to sell essentially the same product (a seat on a flight) to all these different market segments while trying to maximize revenue.

Last, if a given passenger "A" knew that a passenger "B" enjoys the same service as her, on the same flight, on a very similar seat (for instance next to the window) but pays a certain considerable amount lower than her, she would have a negative sentiment towards the company as she would feel, to a certain extent, robbed.

What implications come with this problem? In fact, European law considers dynamic pricing practice as a legitimate procedure that simply captures different customers' willingness to pay. Nevertheless, in a framework where dynamic pricing is used, economic theory shows us that consumer surplus is lower compared to a perfectly competitive scenario with no information asymmetry. Indeed, often many people, for instance students and people travelling for work, have a reservation price that is lower than the price assigned by airline companies but are still buying the ticket because of their tasks as scholars or professionals respectively. As already mentioned above, was the consumer to know how dynamic pricing works, this would entail an issue for the company whose pricing algorithm would be less effective in taking into account the "hostile" and strategic behaviour of aware consumers. However, the mutual effect between the awareness of the consumer and the dynamic pricing strategy is out of the scope of this work and it could be listed as an interesting topic for future works.

3. THEORETICAL MODEL

3.1 Premise: an easy insight of what is meant with dynamic pricing

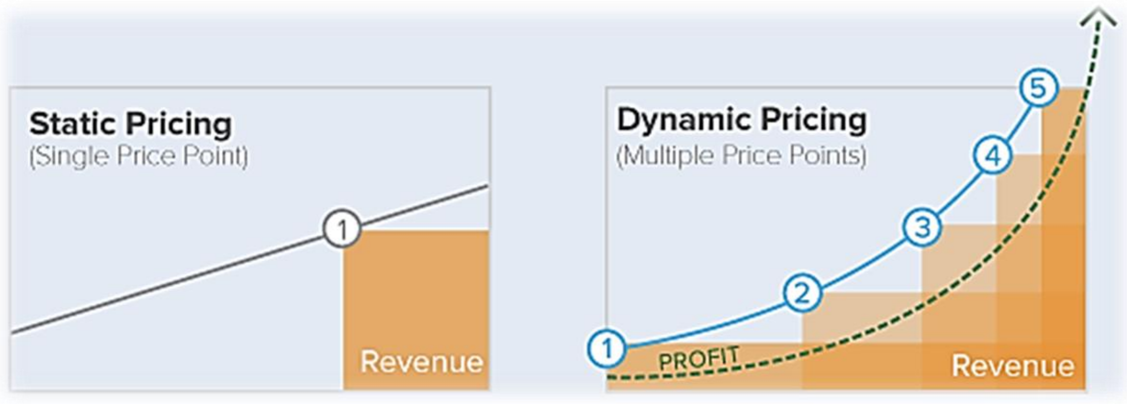

Figure 1. A comparison between static and dynamic pricing. Image extracted from Smyth, "Dynamic Pricing And Price Discrimination: What's The Difference?"

Fixed or static pricing is a strategy in which a price point is established and maintained for an extended period. Dynamic pricing means that the price on a product or service can change over time. As shown in fig. 1, selecting the appropriate strategy for a business has major implications in a company's ongoing effort to attract customers and achieve optimal profit margins.

If we consider the transportation industry, for instance, public bus transportation companies apply static pricing whereas both airlines and train operators prefer to adopt dynamic pricing in order to maximise profits.

Airline pricing both in Europe and in the United States is opaque. It is not uncommon for one-way fares to exceed round-trip prices and to see considerable disruptions in the value of a generic seat from one day to another. The difference in the cabin classes is often a factor to consider but it is not the only one. Prices change frequently, with low fares on a particular flight being available, then not, and then available again. Why is this so? Dynamic Pricing algorithms look at past searches and bookings to see what was offered – airlines, schedules, prices – and what the customer chose. It then balances that data in real-time to find the optimal price. This is the revenue-optimal 'sweet spot' between raising the price to make more money per seat and lowering the price to increase the probability of the customer picking the company's offer. It can be a low price in a competitive market for a price-sensitive customer segment such as tourists. Conversely, it can be a high price if you have a superior product or an impatient customer segment such as executives. Indeed, researches carried out at Massachusetts Institute of Technology (MIT) (Sheppard, *A machine-learning approach to inventory-constrained dynamic pricing*, 2018) show that dynamic pricing can deliver significant revenue gains compared to traditional revenue management alone. In particular, simulations have shown that adding dynamic pricing on top of today's revenue management techniques can lead to revenue gains of up to 3% – 7% when dynamic pricing is used by a single airline in a competitive environment.

To give a more realistic view of how low cost carriers operate in the real world scenario, I will therefore consider an algorithm based on dynamic pricing.

Summary Table of Considered Parameters and Measures

Symbol	**Explanation**
N	Capacity of the plane
T	Total number of booking periods before take-off
t	Observed booking period labelled from T to 0 (assumed to be variable)
M	Total number of seats available at time T
m	Seats available at a given booking period labelled from M to 0
$p\ (t,M)$	Price of a seat when there are t booking periods before takeoff and M seats available
H	Set of consumers from 1 to N who arrives sequentially
$\varphi_{1,t}$	Probability that consumer 1 arrives at time t
$\theta_{h,t}$	Random variable with cumulative F distribution describing the willingness to pay of consumer h at time t
$q\ (p)$	Probability of selling the first available seat at the fare pr
$V\ (t,M)$	Value of the M remaining seats at time t before takeoff
$p^*(t,M)$	Optimal fare for each of the M remaining seats at time t

Table 1: The variables and their corresponding symbols used throughout this work

3.2 Model Assumptions

A company operates a single flight with $N > 1$ seats on a monopolistic route. The flight is sold over $t \geq 1$ booking periods where t is a discrete variable and $t = T, T - 1...2,1$ describes the number of periods remaining before departure. In particular, $t = 1$ is the last booking period while $t = T$ is the first one. For each t, the firm commits to a sequence of fares for all the $M \leq N$ remaining seats of the flight. Thus, until seat $m = M, M\text{-}1...,2,1$ has not been sold, each traveller presenting in booking period t faces a two variables function $p(t,M)$. Within the booking period t, once seat M has been sold, then the next fare on offer becomes $p(t,M-1)$. At the end of the booking period t, the unsold seats are offered in the next period, $t-1$, until t = 1. Seats available at the end of the last booking period remain unsold. In each period t a set of consumers h = 0,1,2,.. arrives sequentially. The probability that the first consumer arrives in t is $\varphi_{1,t} \in (0, 1)$, and that consumer h + 1 arrives conditional on the fact that consumer h has already appeared is $\varphi_{h+1,t} \in (0,1)$. Consumer ($h,t$) is short sighted and her willingness to pay is a random variable $\theta_{h,t}$,with cumulative distribution $F_{h,t}$ on the support $[0, \hat{\theta}_{h,t}]$ with $\hat{\theta}_{h,t} < \infty$. In order to compare the flight's price (p) to the generic consumer h's willingness to pay in time t ($\theta_{h,t}$), p will be a random variable.

Moreover, I make the following simplifying assumptions: for any $h = 0,1,2,..,\infty$ and $t = 1,..,T$ $\varphi_{h,t} = \varphi_{h+1,t} = \varphi \in (0,1)$; $F_{h,t} = F_{h+1,t} = F$, with $\theta_{h,t} = \theta$. Thus, I assume that the process is memoryless (i.e. p is the same for all consumers) and consumers have the same ex-ante evaluation (i.e. the random variable associated to the consumer's willingness to pay is the same for all consumers). The probability of selling the first available seat at the fare p is:

$$q(p) = \varphi\,(1-F(p)) = \sum_{h=0}^{\infty}(\varphi F(p))^h = \frac{\varphi(1-F(p))}{1-\varphi F(p)} \qquad (1)$$

where $\varphi\,(1-F(p))$ is the probability that consumer h arrives and buys at price p provided that consumers 1,..,$h-1$ have previously refused to buy at the same price; and $(\varphi F(p))^h$ is the probability that consumers from 1 to h arrived and did not buy. Not surprisingly, the probability of selling the first available seat is strongly dependent on the price and its cumulative F distribution. Putting on sale the first available seat at a price notably lower than the consumers' expected willingness to pay will increase q and the opposite holds for excessively high prices.

Airline seats are perishable products, that is, there is a finite horizon to sell them, after which any unused capacity is lost. Moreover, the marginal cost of an extra unit of demand is relatively small. For this reason, the models in this work ignore the cost component in the decision-making process and refer to revenue maximization rather than profit maximization.

The maximization problem of the company can be summarized by the following Bellman equation that describes the total value of the M available seats at time t:

$$V(t,M) = \max_p \{q(p)[p + V(t,M-1)] + (1-q(p))V(t-1,M)\} \qquad (2)$$

with boundary conditions $V(t,0) = 0$ and $V(0,M) = 0$, for any $t \in \{0,..,T\}$ and $M \in (0,..,N)$. As (1) showed, $q(p)$ is the probability of selling the first available seat at price p.

Unlike most of the models mentioned in the previous section, the approach in equation (2) assumes the possibility that more than one seat can be sold within each t: this implies the need to set always a different fare for all the seats on an airplane, which is precisely how carriers are believed to operate in real world scenarios. Furthermore, note that equation (2) entails a trade-off between selling now at least one seat (gaining p and the revenue flow coming from the remaining seats, $V(t, M-1)$), and keeping the capacity intact and postpone the sale to the next period, gaining $V(t-1, M)$. This dynamic programming method breaks the decision problem the carrier has to face into smaller subproblems. My aim, however, is to find an optimal strategy for each stage consistent with Bellman's principle of optimality. According to this principle, an optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision. For this reason, I will now focus on finding $p^*(t,M)$, the optimal fare at period t with M seats available.

First order conditions imply that this trade-off is explained by:

$$\Psi(p^*(t, M)) = V(t-1, M) - V(t, M-1), \qquad (3)$$

where $p^*(t,M)$ is the optimal fare when there are t periods and M seats; and $V(t,M)$ is the value of M remaining seats at time t.

$\Psi(p^*) \equiv p^* + q(p^*)/q^1(p^*)$.

Under $\Psi^I > 0$, $p(t, M)$ is unique and can be easily found by inverting (3). Moreover, since $p(t,M)$ only depends on $V(t-1,M)$ and $V(t,M-1)$, the problem described by equation (2) can be easily solved recursively by using equation (3) with the boundary conditions $V(t,0) = 0$ and $V(0,M) = 0$ (see below). This property of the model also implies that $p(t, M)$ is independent of the number of available seats at the start of each t.

<u>Proof</u>

First, note that (2) can be written as:

$$V(t,M) = \max_p \{q(p)[p + V(t,M-1) - V(t-1,M)]\} + V(t-1,M) \qquad (A.1)$$

with boundary conditions $V(t,0) = 0$ and $V(0,M) = 0$, for any $t \in \{0,...,T\}$ and $M \in \{0,...,N\}$. To find a solution for the problem described in (A.1), I consider the following steps.

1. Step 1

   Find the solution for $\max_p q(p)(p + x)$. Since F is bounded a solution for the problem exists. When θ is uniformly distributed in [0,1], there is a closed form solution given by:

$$p = (1 - \sqrt[2]{(1-\varphi)(1+x\varphi)})/\varphi \qquad (A.2)$$

2. Step 2

   To find x, I set $t = 1$ and $M = 1$.

3. Step 3

   Compute $x = V(t,M-1)-V(t-1,M)$, that is, $x = V(1,0) - V(0,1) = 0$ and use Step 1 to get $p(1,1)$. Replace it in (A.1) to obtain $V(1,1)$.

4. Step 4

   Set $M = M + 1$, $x = V(1,1) - V(0,2) = V(1,1)$ that I have already computed previously. Repeat Step 3 until $M = N$.

5. Step 5

   Set $t = t + 1$ and $M=1$. If $t < T$, go back to Step 3.

3.3 Key Propositions

➢ Proposition 1

The value function $V(t, M)$ is increasing in t and M, i.e. $\Delta_1 V(t, M) > 0$ and $\Delta_2 V(t,M) > 0$ where Δ_1 is the forward difference operator with respect to t and Δ_2 is the forward difference operator with respect to M.

<u>Proof</u>

I assume that $\Delta_1 V(t, M)$ is decreasing in t and increasing in M; and $\Delta_2 V(t, M)$ is increasing in t and decreasing in M. Below, I show that these assumptions are satisfied when the willingness-to-pay of travellers is uniformly distributed.

I will show that $V(t,M) > V(t-1,M)$. By contradiction assume that $V(t, M) \leq V(t-1, M)$. Let $p*(t,m)$ with $t = 1,...,t-1$ and $m = 1,...,M$, be the set of fares that solves (1) when there are $t-1$ periods and M seats. Define $\hat{p}=(\tau,m)$ with $\tau = 1,...,t$ and $m = 1,...,M$, as a set of fares (not necessarily the optimal one) that is chosen when there are t periods and M seats: $\hat{p} = (\tau + 1,m) = p*(\tau,m)$, for $\tau = 1,...,t-1$ and $\hat{p}=(1,m) = \bar{p} \in (0, \bar{\theta})$. Then, under this fare profile the expected return gained in the first $t - 1$ periods is $V(t-1, M)$. Because $\varphi < 1$, there is a positive probability that some seats are available in the last period ($t = 1$), and they generate positive expected revenue,which contradicts our assumption. The proof that $V(t, M) > V(t, M-1)$ is similar to the previous case and is omitted.

- Proposition 2

The fare profile {p(t,M),t = 1,...,T;M = 1, ...,N} has the following properties:

1. Invariance: p (t, M) is independent of N.

2. Ascending fare profile: p (t, M) is decreasing in M.

3. Decreasing fares over booking periods: p (t, m) is increasing in t.

Proposition 2 is discussed below and a proof of its statements is given as presented in Alderighi et al. (*The hidden side of dynamic pricing in airline markets*, 2016)

<u>Do statements in Proposition 2 make sense?</u>

Seats	$T=1$	$T=3$			$T=5$				
m	$t=1$	$t=3$	$t=2$	$t=1$	$t=5$	$t=4$	$t=3$	$t=2$	$t=1$
12	0.656	0.691	0.624	0.552	0.703	0.661	0.613	0.563	0.521
11	0.665	0.704	0.636	0.559	0.719	0.676	0.626	0.573	0.526
10	0.676	0.719	0.650	0.567	0.735	0.693	0.642	0.585	0.531
9	0.687	0.735	0.665	0.577	0.752	0.710	0.659	0.598	0.538
8	0.700	0.752	0.681	0.589	0.771	0.730	0.677	0.614	0.547
7	0.714	0.771	0.700	0.602	0.790	0.750	0.699	0.633	0.558
6	0.730	0.791	0.721	0.619	0.811	0.773	0.722	0.655	0.572
5	0.748	0.813	0.745	0.639	0.833	0.798	0.749	0.680	0.590
4	0.770	0.837	0.773	0.664	0.857	0.825	0.779	0.711	0.614
3	0.796	0.864	0.805	0.696	0.883	0.855	0.814	0.749	0.646
2	0.829	0.895	0.844	0.739	0.911	0.889	0.854	0.796	0.692
1	0.875	0.932	0.894	0.805	0.944	0.928	0.904	0.860	0.765

Table 2. Simulated optimal fares p_{jt} in the case of T = 1,3,5 periods. Table extracted from: Alderighi et al., "The hidden side of dynamic pricing in airline markets", page 36

The results of Proposition 2 are illustrated in Table 1, which presents the simulated fares in three different cases: one period (T = 1), three periods (T = 3), and five periods (T = 5). The simulation has been conducted by Alderighi, Gaggero and Piga (*The hidden side of dynamic pricing in airline markets*, 2016). In this case, N, the capacity of the plane, has been set equal to 12, θ, the willingness to pay of customers, is uniformly distributed over [0,1] and φ = {0.9796,0.9412,0.9057} for, respectively, T = {1,3,5}. The values of φ are defined such that the expected number of consumers h is the same in the three cases and equal to 4N = 48. Result 1 of Proposition 2 implies that, conditional on seat m being available, its fare is not affected by the number of seats available on the airplane. Table 1 therefore always reports the fare distribution for all N seats: the Proposition indicates that the optimal fare of, say, seat m = 9 at t = 3 when T = 5 is always 0.659 regardless of whether at t the number of available seats is greater or equal to 9. This result depends on the fact that travelers' arrivals are independent (see above) and therefore, within each period t, only subsequent fares, but not previous fares, if any, affect the optimal level of $p(m,t)$. Moving from the top (first available seat, i.e., seat $m = 12$) to the bottom (last available seat, i.e. seat $m = 1$) of each column, it appears that the fare distribution is increasing both in the one-period and in the multi-period cases. Thus, in any period, consumers who arrive first pay less than those showing up later (Result 2). This is a notable difference from Dana's model where a seller can charge only a single posted price in each period. An ascending fare profile is not novel in the theoretical economic literature, but the explanation proposed here provides interesting extensions I will discuss later on. For instance, in this setup an increasing fare distribution entails that the higher the price, the more unlikely the sales of both current and subsequent seats; that is, a high fare for the seat on sale increases the opportunity cost of having to sell tomorrow all the subsequent seats. The third result in Proposition 2 is illustrated in Table 1 by values of $p(t, m)$ declining over t for any m. This result extends the one-period case considered in Dana (*Equilibrium Price Dispersion under demand uncertainty: the roles of costly capacity and market structure*, 1999) by showing that the carrier's option value decreases as the departure date approaches. This is standard for highly perishable services, as illustrated in Sweeting (*Dynamic Behaviour in Perishable Goods Market*, 2012), where however the analysis is limited to the case of a single ticket and not to a full distribution of prices as in Alderighi and Piga's case. The results in Proposition 2 offer several new empirical implications I will test in the next section of the work. There are however two issues that the theoretical model assumes away: the possibility of strategic consumers and the fact that there is no learning on actual demand during the booking period. In the following part of this Section, I will discuss these potential issues.

3.4 Potential issues to the model

Although my work aims at building a model as close as possible to the real one implemented by Easyjet, there are, among others, two main actual barriers that may hinder the predictive nature of my dynamic pricing algorithm: strategic consumers and unknown demand. Each of these possible limitations will be discussed below.

Strategic Consumers

Proposition 2 shows two contrasting trends as far as the seat on sale is concerned. On the one hand, according to Result 2, within the same period the fare of the next seat is higher than the one on sale. On the other hand, Result 3 states that the fare of a given seat reduces over periods. Thus, the fare of the seat on sale moves up during the same period and down over period switches, especially when the departure date is near. Therefore, the price reductions may be potentially conducive to strategic behavior because the consumer arriving when seat *M* is on sale at time *t* would always prefer to buy it at $t-1$. However, by postponing the purchase, the consumer faces the risk that, at time *t*, other consumers may arrive and buy *M* and some or all subsequent seats. That is, if a consumer expects that the fare of seat $M-1$ will be, on average, higher than that of seat *M*, then strategic behavior is discouraged.

Seats m	Probability of selling, π_{mt} $t=5$	$t=4$	$t=3$	$t=2$	$t=1$	Total	Average paid fare
12	0.747	0.185	0.052	0.014	0.004	1.000	0.688
11	0.567	0.269	0.109	0.046	0.009	0.999	0.689
10	0.421	0.307	0.168	0.075	0.025	0.996	0.690
9	0.301	0.303	0.224	0.120	0.041	0.989	0.691
8	0.202	0.296	0.253	0.162	0.061	0.972	0.694
7	0.129	0.272	0.263	0.193	0.093	0.948	0.699
6	0.085	0.210	0.267	0.227	0.136	0.924	0.703
5	0.052	0.155	0.251	0.249	0.177	0.883	0.711
4	0.028	0.109	0.220	0.253	0.215	0.824	0.724
3	0.015	0.075	0.173	0.249	0.238	0.749	0.744
2	0.008	0.040	0.126	0.236	0.256	0.664	0.774
1	0.002	0.019	0.074	0.180	0.265	0.539	0.822

Table 3: Simulated optimal probabilities of selling (π_{mt}) and average paid fare ($\bar{p}_m$), T = 5 periods. Table extracted from: Alderighi et al., The hidden side of dynamic pricing in airline markets, page 36

In Table 3 a simulation conducted by Alderighi, Gaggero and Piga (*The hidden side of dynamic pricing in airline markets*, 2016) is displayed. Let π *(m,t)* be the probability of selling generic seat m in period *t*, and $\bar{p}$ the average paid fare. Table 2 reports both π *(m,t)* and, in the last column, $\bar{p}$, based on the simulation values of Table 1. Here we notice that the average paid fare is increasing over seats across periods. Thus, consistent with the findings of the simulation presented above, the incentive to postpone a purchase is hindered by the increasing trend of the seat on sale.

Unknown Demand

The demand as a function of price is unknown a priori and is learned over time. Usually carriers set their fares based on three different sources of information: historical data, internal data collected during the booking period, and external data. Historical data are information available to a carrier before its price setting decision. External data are information on demand shifters (e.g., such events as concerts, football matches, etc.) revealed during the booking period. If such information corresponds to an unexpected demand shock, it can be easily accommodated in the model by assuming that a carrier, after receiving it, redesigns a new fare profile based on new values for θ and p. Basically, external data produces a positive or a negative shift of the fare profile from the moment the carrier processes the information onwards. External shocks are by definition very difficult to predict and work with.

4. THE DEVELOPMENT OF A DYNAMIC PROGRAMMING ALGORITHM

4.1 Preliminary Disclosure: a comparison between my theoretical approach and the real world scenario

For the sake of this work, I have considered the assumption that a customer *h,* when deciding to buy an airline ticket at time t, faces a given price *p* as opposed to her willingness to pay, $\theta_{h,t}$. Although being clear and straightforward, this assumption is valid only to a certain extent. For instance, after conducting an experiment consisting of 96 observations, I have noticed that Easyjet customers have to face four different prices simultaneously, which share a common base price, *p*. In particular, to this price, an additive delta is summed and four different classes are obtained. This number depends on both the part of the plane the seats are located and the physical space available for the passenger. The most interesting insight is that while the base price changes over time, these deltas stay fixed and are, therefore, independent from both *M* and *t.*

As these additive constants are fixed and independent from the available seats on the plane and the time interval considered, for the purpose of simplicity as well as to align to the work of Alderighi and Piga, (*The hidden side of dynamic pricing in airline markets*, 2016), I will consider only the simple base price, *p,* a consumer observes at time *t*.

4.2 A first attempt to understand how a dynamic pricing strategy works

In order to understand the functioning of the dynamic pricing strategy low cost airline companies follow, I have conducted an experiment of 96 observation in six days, on different Easyjet flights on the Rome-Amsterdam route. Most conspicuously, the query dates were set such that flights entered my database respectively one week and one month before departure; then they were surveyed at 6-hours distanced intervals for 6 day, to get a better understanding of the price evolution as the date of departure neared. The website response to the query included flight information, for each observation, for the two dates: the base price, the additive deltas identifying four different classes and the number of seats purchased for each class at the given posted fare. This information is important to derive the price distributions from the posted fares. Nevertheless, by assuming four different prices observed by the consumer and the fact that seats are not perfectly identical, capturing the essence of how dynamic pricing works may be difficult and out of the scope of this work. Moreover, it may be noticed that these additive deltas are constant over time and do not change as seats in each class are sold. Thus, as mentioned above, I will focus on simply one price when moving to further approaches.

4.3 A second approach: the retrieval of information using a crawler

As it should be clear by now, customers are currently facing new challenges and need to learn how to process all available market data, gain useful insights, and evaluate outcomes. The easiest way to achieve this is by having a dynamic pricing strategy that uses automated techniques. For this reason, a second way to understand how dynamic pricing works was to develop and use a crawler. A web crawler is a bot that searches for a set of information. It goes through the website, and finds the required data or keywords that were mentioned as a search topic. Each crawler is different, but what normally they have in common is that they go inside a website and then start going deeper and deeper on the links and pages, scraping every page and saving its content.

The application of crawlers in the airline industry is not new, as crawlers have been widely used in studying how revenue management works. For instance, Li et al. (*Are Consumers Strategic? Structural Estimation from the Air-Travel Industry*, 2014), assuming different levels of sophistication in consumers' perception of future prices, estimated the fraction of strategic consumers in the population, in order to maximise revenues for airline companies. Indeed, ideally, all the fares of a given flight and their trend can be obtained from the internet, using a web spider, which accessed the websites of low cost carriers and retrieved all the information requested. This approach is very interesting as most of the empirical literature on airline pricing focuses on the price of one seat, that corresponding to the seat being on sale at the time of the query. A central contribution of recent works such as the one of Alderighi and Piga (*The hidden side of dynamic pricing in airline markets*, 2016) is to show that this is not sufficient to test the implications of theoretical models of DP in airline markets. Based on the model presented in Section 3.2, my ideal data collection incorporated an experimental design explicitly aimed at recovering a flight's price distribution, as it was actually stored on the carriers' web reservation system.

In practice, however, this was not possible, at least on Easyjet's website. Indeed, Easyjet employs a reservation page using dynamic HTML language. In other words, the information on pricing for instance is not always available but it is uploaded dynamically as the customer proceeds with her reservation. This issue is daunting for the website user interested in the algorithm adopted by Easyjet as she has to develop a customised crawler for the website. Most conspicuously, the above-mentioned crawler should be able to analyse the HTML code of the page, and formulate the query accordingly to retrieve data on the route and the price. Due to this issue, I have moved to the third approach: a manual enquiry.

4.4 The third approach: an experimental design to recover Easyjet price distribution

As already mentioned, most of the empirical literature on airline pricing focuses on the price of one seat, that corresponding to the seat being on sale at the time of the query. However, a central contribution of the paper of Alderighi and Piga (*The hidden side of dynamic pricing in airline markets*, 2016), which is used as a reference point in the present work, is to show that this is not sufficient to test the implications of theoretical models of DP in airline markets.

Indeed, based on the model presented in Sections 3.2 and 3.3, I have set up a data collection which incorporates an experimental design explicitly aimed at recovering a flight's price distribution, as it is stored on the Easyjet's web reservation system. Actually, this implied the implementation of the following procedure. For flights with different departure dates, one week and one month, I manually started by requesting the price of one seat, and then continued by sequentially increasing the number of seats by one unit. The sequence would stop either because the maximum number of seats in a query, equal to 40, was reached or at a smaller number of seats, depending on the number of seats available for the flight considered. Theoretically, if we used a crawler, as in Alderighi et al. (2015), the latter case would directly indicate the exact number of seats available on the flight on the particular query date and we could store in a variable called "Available Seats" in order to track how a flight occupancy changes as the departure date nears. On the other hand, the former case would correspond to a situation where is known that at least 40 seats remain to be sold on a given query date. In other words, as it is should be clear by now, the number of available seats purchasable by a consumer for a given query on Easyjet website is limited by the system to 40.

Consistent with the findings of Alderighi et al. (*Targeting leisure and business passengers with unsegmented pricing*, 2015), I expect to find a U-shaped distribution of posted fares driven by the presence of *C*, a fixed charge or "commission per booking". In particular, Alderighi et al, (*Targeting leisure and business passengers with unsegmented pricing*, 2015) through data visual inspection, learnt that the carriers' posted fare follow this rule:

$$PF(s) = \frac{C + \sum_{j=1}^{s} p_j}{s} \tag{4}$$

where *s* denotes the number of seats in the query, *PF(s)* the corresponding posted fare, p_j the fare of each seat, starting from the first one available for sale. As already mentioned, the presence of *C* implies that the distribution of posted fares over seats is generally U-shaped, as the commission is spread over more seats while the ratio between the sum of prices and the number of seats in a query increases over time due to the increasing values of the buckets.

To succeed in reverse-engineering the carrier's pricing approach, I will start by determining *C*. To find *C*, I rely on the fact that in most cases the first and the second seat are likely to belong to the same bucket (Gaggero & Piga, *Airline Competition in the British Isles*, 2010). Therefore *C* (and the value of the first bucket) can be obtained by solving the following system of two linear equations in two unknowns, using the identity $p_1 = p_2 = p$:

$$PF(1) = C + p \tag{5}$$

$$PF(2) = \frac{C + 2p}{2} \tag{6}$$

After finding *C*, using (6) and (7) it is straightforward to derive the bucket fare tags, P_j

$$Pj = j * PF(j) - (j-1) * PF(j-1) \text{ with } j \in [2,40] \tag{7}$$

with $P_1 = PF(1) - C$.

I will now compute *C* on the observation for *t* = 1 and *m*=1 for Easyjet flight Rome-Amsterdam with departure date in one month. The prices for different t and m can be found in Section 5.

Proof of Concept

$PF(1) = €80{,}19$ *and* $PF(2) = €70{,}19$

$C = PF(1) - p$ *from* (5) (B.1)

$PF(2) = \frac{PF(1) + p}{2}$ *from* (B.1) *and* (6) (B.2)

$p = \underline{€60{,}\ 19}$ *and* $C = €20{,}00$

$P_2 = 2*PF(2) - PF(1) = \underline{60{,}19}$

Not surprisingly, the two values coincide.

In an attempt to understand how Easyjet implements its proprietary pricing algorithm, a second step in my analysis consists in developing a programme using Python that resembles the functioning of the real algorithm used by the company, as in Alderighi, Gaggero and Piga (*The hidden side of dynamic pricing in airline markets*, 2016). As in that reference, the values predicted by the algorithm have to be normalized through a normalizing factor that enables the user to compare her forecasts to actual prices. Indeed, the algorithm that, by construction, starts by setting its prices to a standard value (e.g. $p = 0$) has to be initialled to a reasonable value as to allow its correct functioning. Was the programme set in the wrong way, it would lose much of its predicting power forecasting unrealistic values from an empirical point of view (i.e. prices that are too high or too low). My algorithm gives as an output a price matrix row by column 27 x 20 corresponding to 27 observations in different time intervals of 20 seats of a given flight. However, a word of caveat is given: contrary to Alderighi, Gaggero and Piga, I have focused on just one week of evidence, as opposed to 4 month-time, and this may have given just a partial insight of the whole fare distribution. To compensate for this scenario diversity I have considered appropriate intertemporal normalizing variables as to have a similar - and thus comparable - starting condition. In order to understand to what extent my programme was a reliable proxy of the pricing algorithm used by the carrier, I have collected a series of 27 observations for two different Easyjet's flights in the Rome-Amsterdam route with different departure dates, respectively one week and one month. The results have been particularly interesting and are described thoroughly 5.2. Through a statistical analysis on the quartiles, I have observed a systemic error of at least 20% when comparing my predictions to actual prices. In considering the flight whose departure date is in one-month time, the error can increase up to around 25%. A possible explanation is that the normalizing factor used to initialize my algorithm is too high and this negatively affects the accuracy of its predicting power, especially in the long term perspective. Furthermore, it must be considered that my data have been retrieved in May so, especially for the latter flight, a seasonality variable may have blurred the results. Perhaps, for future research, other time grids could do a better job.

It could be argued that this is a peculiar approach followed by the carrier considered in my experiment. However, the empirical findings in Alderighi et al. (*Targeting leisure and business passengers with unsegmented pricing*, 2015) also suggest that another European low cost carrier, Ryanair, defines a similar fare distribution across seats. As far as so-called "traditional" legacy carriers are concerned, the analysis is complicated by their adoption of a nested-classes system, where the same seat can belong to different classes, each with different ticket restrictions. Nonetheless, various papers present graphical evidence of fares whose temporal path also follows a step-wise pattern, with each step representing a class, i.e.,"bucket", level (Lazarev, *The Welfare Effects of Intertemporal Price Discrimination: An Empirical Analysis of Airline Pricing in U.S. Monopoly Markets*, 2013; McAfee and te Velde, *Dynamic Pricing in the Airline Industry*, 2007; Puller et al., *Testing theories of Scarcity pricing and Price Dispersion in the Airline Industry*, 2009).

4.5 Does a fixed probability of arrival make sense?

Arguably, I have considered a fixed probability of arrival for the development of my dynamic pricing algorithm. While this assumption is reasonable in order to test and verify my approach, a model that better captures a real world situation should consider a variating trend for φ. A possible trend could be the one followed by Alderighi and Piga, (*The hidden side of dynamic pricing in airline markets*, 2016) that is a decreasing trend as the departure date approaches. Indeed, the two consider three different probability of arrival, φ ={0.9796, 0.9412, 0.9057}, one for each time period considered, T = {1, 3, 5}. Another interesting point of view is a variating Gaussian trend as opposed to a monotonic one. In other words, at first, the probability of arrival is relatively low and then increases over time. For instance, it could be that, for a given flight, one week before the take-off, φ reaches its peak, after which it decreases again. Airline companies know this trend in advance from their historical data and, thus, are able to anticipate consumer behaviour and maximise their revenues. With respect to the probability of arrival and its trend, future models should keep the impact of this assumption into account, while consider also potential trade-offs between simplicity and realism.

5. DESCRIPTIVE STATISTICS AND DATA ANALYSIS

5.1 Disclosure on the functioning of my pricing algorithm in Python

Actual revenue management practices are usually not limited to developing and using a revenue maximizing algorithm yet they entail solving almost in real time several operations research problems, designed to constantly try to match supply and demand while, at the same time, ideally perfectly capturing consumer surplus. In other words, the actual revenue maximization problem requires many complex mathematical tools that are out the scope of this work. However, by applying my knowledge in Python, I have been able to build a relatively simple dynamic pricing algorithm whose functioning resembles the one of Alderighi and Piga (*The hidden side of dynamic pricing in airline markets,* 2016), presented in 3.2. The outcome is a *t x m* matrix that shows at a given time the fare distribution of a flight based on the booking periods left before takeoff as well as the available seats on the plane.

In the following paragraph, I will briefly explain what Python is and how my algorithm works. Python is a programming language first introduced in 1991. Python features a dynamic type system and automatic memory management. Furthermore, it supports multiple programming paradigms and has a large and comprehensive standard library. As for my algorithm, I started by importing the appropriate libraries for the scope of the work: math and random. Math provides access to most mathematical functions and random generates pseudorandom numbers with arbitrary precision decimals. Next, I have defined four functions: *InitV, F, Q* and *ScriviMatrice. InitV* is a function that takes *T* time intervals and *M* seats and puts them in *T* x *M* a "list of lists"(i.e. the future matrix). *F* is a function that, in order to study the F distribution of the price and the willingness to pay theta, takes these two parameters and assesses whether the price is greater, lower or equal to theta as to retrieve their distribution. *Q* is a function that resembles the variable *q* defined in 3.2 while *ScriviMatrice* exports on a comma separated values (CSV) file a matrix, and then "cleans" it so that the data can be easily displayed and analysed on other platforms such as Excel or R. A screenshot of the IDLE display presenting these functions is found below.

```
import math
import random

def initV(T,M):
    v=[]
    for i in range(0,T+1):
        v.append([])
        for j in range(0,M+1):
            v[i].append(0)
    return v

def F(p,th):
    if p <= 0:
        return 0
    elif p <= th:
        return p/th
    elif p> th:
        return 1

def Q(price,fi,t):
    num=fi*(1-F(price,t))
    den=1-fi*F(price,t)
    result=num/den
    return result

def Scrivimatrice(Mat):
    fout= open("Matrice.csv","w", encoding="UTF-8")
    for i in range(0,T):
        for j in range (0,M):
            stringa=str(Mat[i][j])
            valore=stringa.replace(".",",")
            #print(valore)
            fout.write(valore+";")
        print(file=fout, end="\n")
    fout.close()
    return
```

Figure 2: IDLE screen presenting the functions used by my algorithm

In the second stage, I wrote the code taking into account the functions defined before. I have hypothesized a two possible worlds scenario: the first, more similar to the real case, supposes the probability of arrival to be variable over time while the latter, consistent with the theoretical model, assumes it is fixed. I have set this variable, *realcase,* to *False*, that is, I have assumed the probability of arrival to be fixed. In this way, future research could just change this variable and set *p*, the price, accordingly. The probability of arrival has been set to 0.9412, consistent with the experiments conducted by Alderighi and Piga (*The hidden side of dynamic pricing in airline markets*, 2016) and, as described in 3.2, *m* and *t* have been set initially to 1. By analysing several Easyjet flights on the Rome-Amsterdam route, which will be the object of my observation as presented in 5.2, I have assessed that the average capacity of a plane is 132 seats, thus *M* has been set to this value. For the sake of simplicity and to align my algorithm to the experiment of the following section, I have assumed only 28 time periods, *t* which are initially set to *T* and then decrease over time (i.e. *T-1, T -2..0*), as described in 3.2. Again in 3.2, it is also explained what *x* is and why both this variable and *p* are initially set to 0. Furthermore, the code is made up of two for cycles which assess the price of every seat for each booking period, consistent with the actual revenue management procedure described in 2.2. Finally, through the function *ScriviMatrice*, the different values of *p* for every seat in each booking period are then inserted in a CSV file and imported in Excel. In the pictures below, the IDLE screen with the second part of the code is presented and the printing of the values in the Python format shown.

File Edit Format Run Options Window Help

```
    for i in range(0,T):
        for j in range (0,M):
            stringa=str(Mat[i][j])
            valore=stringa.replace(".",",")
            #print(valore)
            fout.write(valore+";")
        print(file=fout, end="\n")
    fout.close()
    return


realcase=False
listPrices=[]
poa = 0.9412 #probability Of Arrival (fi)
t=1
m=1
T=28 #max number of time instants
M=132 #max number of seats
V=initV(T,M) #matrix containing all the values
x=0
p=0
for t in range(1,T+1):
    theta=random.uniform(0,1)
    #poa=poa-0.05
    for m in range(1,M+1):
        x=V[t][m-1] - V[t-1][m]
        if (realcase) and (t==1) and (m==1):
            p=136
        else:
            p=(1-math.sqrt((1-poa)*(1+x*poa)))/poa
        listPrices.append(p)
        #compute v(t,m)
        V[t][m]=Q(p,poa,theta)*(p+x)+V[t-1][m]
        print(t,m,p)
Scrivimatrice(V)
```

Figure 3: IDLE screen presenting the second part of the pricing algorithm

```
28 64 0.832028155633404
28 65 0.8312143631952674
28 66 0.8304039302358756
28 67 0.829598781651315
28 68 0.8288006802982921
28 69 0.8280112257594755
28 70 0.827231856249588
28 71 0.826463852959771
28 72 0.8257083462244177
28 73 0.8249663229848995
28 74 0.8242386351122232
28 75 0.8235260082318177
28 76 0.822829050766236
28 77 0.8221482629747129
28 78 0.8214840458221966
28 79 0.8208367095552127
28 80 0.8202064818984922
28 81 0.8195935158157239
28 82 0.8189978968010104
28 83 0.8184196496856255
28 84 0.8178587449583583
28 85 0.8173151046078438
28 86 0.8167886075025818
28 87 0.8162790943293313
28 88 0.8157863721137948
28 89 0.8153102183493472
```

Figure 4: Python screen presenting the values printed by my algorithm. The first column corresponds to t, the second to m and the third to the price before normalization.

5.2 My experiment: an assessment of the efficiency of my pricing algorithm

In this Section, I will present the results of a pricing simulation conducted through my algorithm as compared to the actual prices charged by Easyjet under a real world scenario: a plane leaving in one week in the Rome–Amsterdam route in the month of April. As mentioned in 5.1, my pricing algorithm resembles the crawler used by Alderighi and Piga (*The hidden side of dynamic pricing in airline markets*, 2016) yet, for the sake of simplicity, I consider my experiment to take place in only 27 periods of observations, *t*, represented in the first column. In order to determine to what extent my algorithm is able to predict the actual fare distribution of Easyjet's flights, I study the how fares for the first available seat evolve over time for the *m* seats available at time t. In my matrix, the second column indicates the number of seats available. This number is possible to be determined anytime by the user as the Easyjet website, after the route and the number of tickets have been inserted, displays the capacity of the plane as well as the number of seats already bought. The third column shows the actual price observed, *p*, after deducting the booking cost estimated in 4.4 while the fourth column displays the plain price, *k*, estimated by my algorithm without any booking cost. It is useful to consider that Easyjet presents four different "classes" of tickets for its flight: the price displayed in the third column is just the “basic” class with no additional services. The values have been inserted according to the Italian convention, with “,” instead of “.” as decimal separator.

Booking period	Available Seats	Actual price	Estimated Price
27	67	€ 60,00	€ 54,88
26	67	€ 60,00	€ 60, 53
25	66	€ 60,00	€ 58,61
24	66	€ 50,91	€ 60, 48
23	63	€ 50,91	€ 43,02
22	63	€ 71,61	€ 83,34
21	51	€ 71,61	€ 64,17
20	48	€ 85,75	€ 67,94
19	46	€ 85,75	€ 83,42
18	43	€ 85,75	€ 83,52
17	36	€ 85,75	€ 84,34
16	32	€ 85,75	€ 71,27
15	26	€ 85,75	€ 85,21
14	23	€ 85,75	€ 72,30
13	21	€ 85,75	€ 85,30
12	16	€ 85,75	€ 86,61
11	13	€ 85,75	€ 77,84
10	11	€ 85,75	€ 87,12
9	11	€ 85,75	€ 81,85
8	11	€ 85,75	€ 73,17
7	8	€ 85,75	€ 84,85
6	8	€ 85,75	€ 85,00
5	7	€ 85,75	€ 81,14
4	7	€ 85,75	€ 82,84
3	7	€ 85,75	€ 82,84
2	4	€ 85,75	€ 84,35
1	1	€ 85,75	€ 84,48

Table 4 shows a comparison between the actual and estimated price at time t for m available seats

First, as for the second column, it is possible to notice that the uniform probability of purchase stated in 3.2 does not hold. Indeed, buyers notably vary the amount of tickets bought from one booking period to the other. Second, it is interesting to notice that actual prices tend to follow a u-shape, consistently with the findings of Alderighi and Piga (*The hidden side of dynamic pricing in airline* *markets*, 2016) presented in 4.4. Third, my algorithm, after normalizing its prices through an appropriate normalizing factor, is able to predict with a considerable degree of precision the actual prices charged by Easyjet. Nonetheless, it is not able to rebuild the actual u-shape distribution the actual prices tend to follow. Last, a word of caveat. For the sake of simplicity, I have considered only the price for the first available seat at time *t*. It must be noted that this is done just to measure the efficiency of my pricing algorithm. I have claimed that focusing the empirical analysis on the fare for the first seat on sale is not a valid way to conduct a test on the bucket intertemporal distribution for the reasons discussed in 2.1. Future research should try to assess the prices for all the available seats at a given time.

5.3 Analysis of the bucket fare distribution under two real world scenarios

In this Section, I will focus on showing the existence of the buckets and on determining their evolution over time under two real world scenarios on the Easyjet Rome-Amsterdam route: I will study the bucket distribution for a plane leaving in one week and one leaving in one month. Due to the relatively low amount of data, the experiment follows a manual data collection approach for both flights and is constituted by 27 time periods. The observations took place every six hours for a week between the months of April and May. For this analysis, I will introduce the variable *x* that represents the number of tickets bought through each query on the Easyjet website and it is by default initialized to the value *X*. In both the experiments, *X* has been set equal to 20. While the choice to consider the time interval between the observations for the two planes to be the same may not seem appropriate from an empirical point of view, my experiment is still consistent with the approach followed by Alderighi and Piga (*The hidden side of dynamic pricing in airline markets*, 2016). Indeed, to the extent that the total number of booking periods observed, *T*, is greater than or equal to the total number of seats initially available, *M*, - as to capture any possible variation in the bucket size and price due to a given sale - both any number *T* and time interval chosen are reasonable. In particular, in their experiment Alderighi et al. surveyed different flights at 10-days distanced intervals until 30 days, and then at more frequent intervals (21, 14, 10, 7, 4 and 1) to get a better understanding of the price evolution as the date of departure neared. Naturally, from an empirical point of view, the probability of selling a ticket for a plane departing in one week is higher than the one for a plane departing in one month and this alone would entail considering different time intervals accordingly. Nevertheless, for the reasons mentioned above, my approach is reasonable and considerations on the two flights can be easily drawn.

The outcome for the flight leaving in one week, under the assumptions that $X = 20$ and $T = 27$, is found below. For graphical reasons, the table has been divided in two parts corresponding to the two halves of the matrix, cut vertically between column x = 11 and x = 10. Again, the values have been inserted according to the Italian convention, with "," instead of "." as decimal separator.

	x = 20	x = 19	x = 18	x = 17	x = 16	x = 15	x = 14	x = 13	x = 12	x = 11
t = 27	€ 60,19	€ 60,19	€ 60,20	€ 60,18	€ 71,79	€ 71,85	€ 71,81	€ 71,79	€ 71,83	€ 71,77
t = 26	€ 60,19	€ 60,19	€ 60,20	€ 60,18	€ 71,79	€ 71,85	€ 71,81	€ 71,79	€ 71,83	€ 71,77
t = 25	€ 60,19	€ 60,19	€ 60,20	€ 60,18	€ 71,79	€ 71,85	€ 71,81	€ 71,79	€ 71,83	€ 71,77
t = 24	€ 51,10	€ 51,10	€ 51,11	€ 51,09	€ 60,20	€ 60,22	€ 60,20	€ 60,18	€ 60,22	€ 60,08
t = 23	€ 51,10	€ 51,10	€ 51,11	€ 51,09	€ 60,20	€ 60,22	€ 60,20	€ 60,18	€ 60,22	€ 60,08
t = 22	€ 71,80	€ 71,80	€ 71,81	€ 71,79	€ 71,80	€ 86,00	€ 85,90	€ 85,90	€ 85,99	€ 85,91
t = 21	€ 71,80	€ 85,94	€ 85,96	€ 85,94	€ 85,91	€ 85,97	€ 85,94	€ 85,90	€ 86,04	€ 85,90
t = 20	€ 85,94	€ 85,94	€ 85,95	€ 85,93	€ 85,94	€ 85,98	€ 85,92	€ 85,92	€ 86,01	€ 99,47
t = 19	€ 85,94	€ 85,94	€ 84,60	€ 87,28	€ 85,94	€ 85,98	€ 85,92	€ 85,92	€ 86,01	€ 99,47
t = 18	€ 85,94	€ 85,94	€ 85,95	€ 85,93	€ 85,94	€ 85,98	€ 85,92	€ 99,60	€ 99,60	€ 99,50
t = 17	€ 85,94	€ 85,94	€ 85,95	€ 85,93	€ 85,94	€ 85,98	€ 85,92	€ 99,60	€ 99,60	€ 99,50
t = 16	€ 85,94	€ 85,94	€ 85,95	€ 85,93	€ 85,94	€ 85,98	€ 99,57	€ 99,55	€ 99,68	€ 99,52
t = 15	€ 85,94	€ 85,94	€ 85,95	€ 85,93	€ 85,94	€ 85,98	€ 99,57	€ 99,55	€ 99,68	€ 99,52
t = 14	€ 85,94	€ 85,94	€ 85,95	€ 85,93	€ 85,94	€ 85,98	€ 99,57	€ 99,55	€ 99,68	€ 99,52
t = 13	€ 85,94	€ 85,94	€ 85,95	€ 85,93	€ 85,94	€ 85,98	€ 99,57	€ 99,55	€ 99,68	€ 99,52
t = 12	€ 85,94	€ 85,94	€ 85,95	€ 85,93	€ 99,59	€ 99,63	€ 99,57	€ 99,53	€ 99,67	€ 99,45
t = 11	€ 85,94	€ 85,94	€ 85,95	€ 85,93	€ 99,59	€ 99,63	€ 99,57	€ 99,53	€ 99,67	€ 99,45
t = 10	€ 85,94	€ 85,94	€ 85,95	€ 85,93	€ 99,59	€ 99,63	€ 99,57	€ 99,53	€ 99,67	€ 99,45
t = 9	€ 85,94	€ 85,94	€ 85,95	€ 99,57	€ 99,60	€ 99,60	€ 99,53	€ 99,63	€ 99,58	€ 99,56
t = 8	€ 85,94	€ 85,94	€ 85,95	€ 99,57	€ 99,60	€ 99,60	€ 99,53	€ 99,63	€ 99,58	€ 99,56
t = 7	€ 85,94	€ 85,94	€ 85,95	€ 99,57	€ 99,60	€ 99,60	€ 99,53	€ 99,63	€ 99,58	€ 99,56
t = 6	€ 85,94	€ 85,94	€ 85,95	€ 99,57	€ 99,60	€ 99,60	€ 99,53	€ 99,63	€ 99,58	€ 99,56
t = 5	€ 85,94	€ 85,94	€ 85,95	€ 99,57	€ 99,60	€ 99,60	€ 99,53	€ 99,63	€ 99,58	€ 99,56
t = 4	€ 85,94	€ 85,94	€ 85,95	€ 99,57	€ 99,60	€ 99,60	€ 99,53	€ 99,63	€ 99,58	€ 99,56
t = 3	€ 85,94	€ 85,94	€ 85,95	€ 99,57	€ 99,60	€ 99,60	€ 99,53	€ 99,63	€ 99,58	€ 99,56
t = 2	€ 85,94	€ 85,94	€ 85,95	€ 99,57	€ 99,60	€ 99,60	€ 99,53	€ 99,63	€ 99,58	€ 99,56
t = 1	€ 85,94	€ 85,94	€ 85,95	€ 99,57	€ 99,60	€ 99,60	€ 99,53	€ 99,63	€ 99,58	€ 99,56

Table 5 : Actual prices for M= 20 seats collected for the Easyjet Rome – Amsterdam flight departing in one week.

	x = 10	x = 9	x = 8	x = 7	x = 6	x = 5	x = 4	x = 3	x = 2	x = 1
t = 27	€ 71,80	€ 71,80	€ 71,81	€ 85,87	€ 86,07	€ 85,85	€ 86,06	€ 86,00	€ 85,86	€ 85,88
t = 26	€ 71,80	€ 71,80	€ 71,81	€ 85,87	€ 86,07	€ 85,85	€ 86,06	€ 86,00	€ 85,86	€ 85,88
t = 25	€ 71,80	€ 71,80	€ 71,81	€ 85,87	€ 86,07	€ 85,85	€ 86,06	€ 86,00	€ 85,86	€ 85,88
t = 24	€ 71,86	€ 71,88	€ 71,69	€ 71,83	€ 71,94	€ 71,62	€ 85,96	€ 86,08	€ 85,97	€ 85,87
t = 23	€ 71,86	€ 71,88	€ 71,69	€ 71,83	€ 71,94	€ 71,62	€ 85,96	€ 86,08	€ 85,97	€ 85,87
t = 22	€ 85,93	€ 86,01	€ 85,88	€ 85,96	€ 99,72	€ 99,48	€ 99,61	€ 99,59	€ 99,65	€ 99,47
t = 21	€ 99,51	€ 99,71	€ 99,50	€ 99,62	€ 99,56	€ 99,52	€ 99,58	€ 99,82	€ 118,18	€ 118,10
t = 20	€ 99,64	€ 99,60	€ 99,60	€ 99,52	€ 99,74	€ 99,46	€ 99,62	€ 99,66	€ 99,58	€ 99,38
t = 19	€ 99,64	€ 99,60	€ 99,60	€ 99,52	€ 99,74	€ 99,46	€ 99,62	€ 99,66	€ 99,58	€ 99,38
t = 18	€ 99,62	€ 99,56	€ 99,66	€ 99,52	€ 99,59	€ 99,51	€ 99,62	€ 99,82	€ 99,39	€ 118,21
t = 17	€ 99,62	€ 99,56	€ 99,66	€ 99,52	€ 99,59	€ 99,51	€ 99,62	€ 99,82	€ 99,39	€ 118,21
t = 16	€ 99,56	€ 99,60	€ 99,50	€ 99,58	€ 99,66	€ 99,62	€ 99,63	€ 99,55	€ 118,38	€ 118,12
t = 15	€ 99,56	€ 99,60	€ 99,50	€ 99,58	€ 99,66	€ 99,62	€ 99,63	€ 99,55	€ 118,38	€ 118,12
t = 14	€ 99,56	€ 99,60	€ 99,50	€ 99,58	€ 99,66	€ 99,62	€ 99,63	€ 99,55	€ 118,38	€ 118,12
t = 13	€ 99,56	€ 99,60	€ 99,50	€ 99,58	€ 99,66	€ 99,62	€ 99,63	€ 99,55	€ 118,38	€ 118,12
t = 12	€ 99,64	€ 99,56	€ 99,56	€ 99,58	€ 99,66	€ 99,52	€ 118,33	€ 118,27	€ 118,36	€ 118,12
t = 11	€ 99,64	€ 99,56	€ 99,56	€ 99,58	€ 99,66	€ 99,52	€ 118,33	€ 118,27	€ 118,36	€ 118,12
t = 10	€ 99,64	€ 99,56	€ 99,56	€ 99,58	€ 99,66	€ 99,52	€ 118,33	€ 118,27	€ 118,36	€ 118,12
t = 9	€ 99,58	€ 99,60	€ 99,53	€ 99,65	€ 99,59	€ 118,19	€ 118,31	€ 118,37	€ 118,26	€ 143,82
t = 8	€ 99,58	€ 99,60	€ 99,53	€ 99,65	€ 99,59	€ 118,19	€ 118,31	€ 118,37	€ 118,26	€ 143,82
t = 7	€ 99,58	€ 99,60	€ 99,53	€ 99,65	€ 99,59	€ 118,19	€ 118,31	€ 118,37	€ 118,26	€ 143,82
t = 6	€ 99,58	€ 99,60	€ 99,53	€ 99,65	€ 99,59	€ 118,19	€ 118,31	€ 118,37	€ 118,26	€ 143,82
t = 5	€ 99,58	€ 99,60	€ 99,53	€ 99,65	€ 99,59	€ 118,19	€ 118,31	€ 118,37	€ 118,26	€ 143,82
t = 4	€ 99,58	€ 99,60	€ 99,53	€ 99,65	€ 99,59	€ 118,19	€ 118,31	€ 118,37	€ 118,26	€ 143,82
t = 3	€ 99,58	€ 99,60	€ 99,53	€ 99,65	€ 99,59	€ 118,19	€ 118,31	€ 118,37	€ 118,26	€ 143,82
t = 2	€ 99,58	€ 99,60	€ 99,53	€ 99,65	€ 99,59	€ 118,19	€ 118,31	€ 118,37	€ 118,26	€ 143,82
t = 1	€ 99,58	€ 99,60	€ 99,53	€ 99,65	€ 99,59	€ 118,19	€ 118,31	€ 118,37	€ 118,26	€ 143,82

Table 5 bis: Actual prices for M = 20 seats collected for the Easyjet Rome – Amsterdam flight departing in one week.

Below, the main descriptive statistics are shown.

Mean		€ 94,03
Median		€ 99,53
St.Deviation		16,05
Variance		257,56
Min		€ 51,09
Max		€ 143,82
Range		€ 92,73
1st Quartile		€ 85,94
3rd Quartile		€ 99,62

Table 6: Descriptive Statistics for the Easyjet Roma-Amsterdam flight departing in one week.

Looking at the descriptive statistics, some interesting evidences can be be drawn. Contrary to my predictions, the statistical distribution of the fares is skewed to the right. Moreover, surprisingly, the minimum fare does not correspond to the first observation. However, this is consistent with the findings of Alderighi and Piga (*The hidden side of dynamic pricing in airline markets*, 2016) and revenue management theory and indicates that the carrier is pursuing a pricing strategy such to attract more customers by reducing the price some days before the take-off. On the other hand, the maximum price is observed during the latest observations as to indicate that the company is targeting specifically business travellers while discouraging strategic buyers who have postponed the purchase of their tickets hoping for some last-minute discount. Finally, note that the standard deviation is considerably high as to denote the not negligible amount of variation in the set of prices possibly due, among other variables, to exogenous factors (e.g. seasonality).

Below graphs for t = 27, t = 18, t = 10 and t =1 can be found. On the x-axis, I consider the number of tickets for which the query has been made. On the other hand, on the y-axis, I plot the price as a function of x.

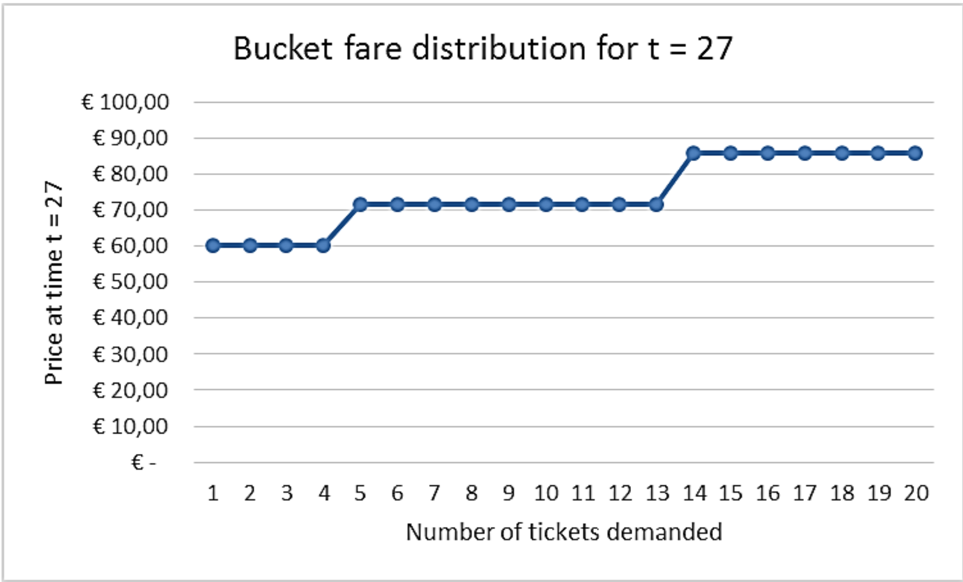

Figure 5: bucket fare distribution for t = 27 for the flight leaving in one week

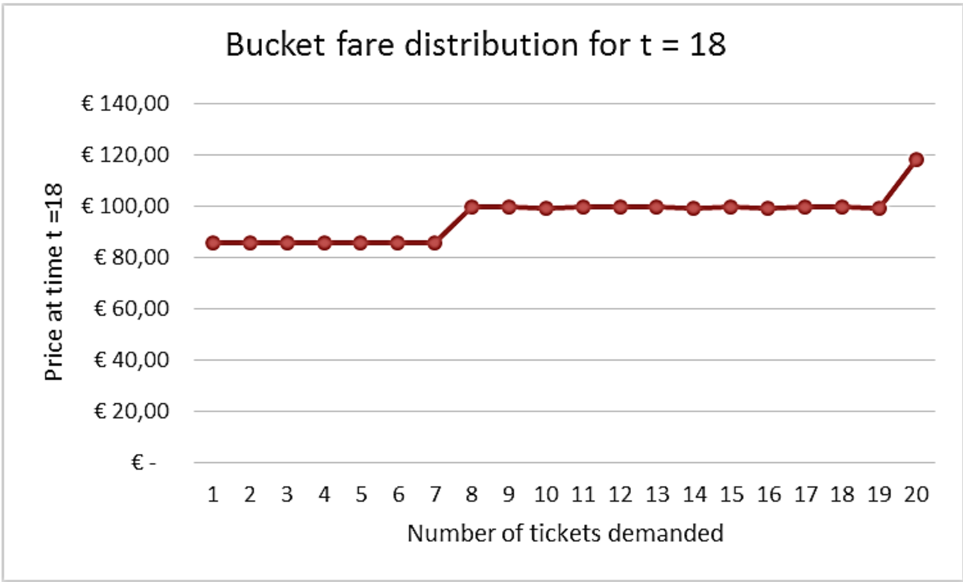

Figure 6: bucket fare distribution for t = 18 for the flight leaving in one week

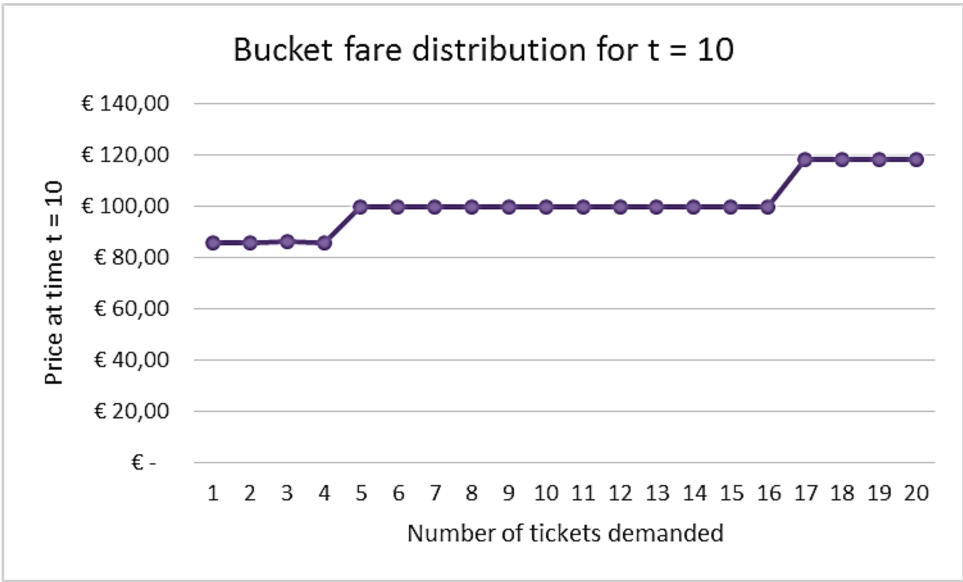

Figure 7: bucket fare distribution for t = 10 for the flight leaving in one week

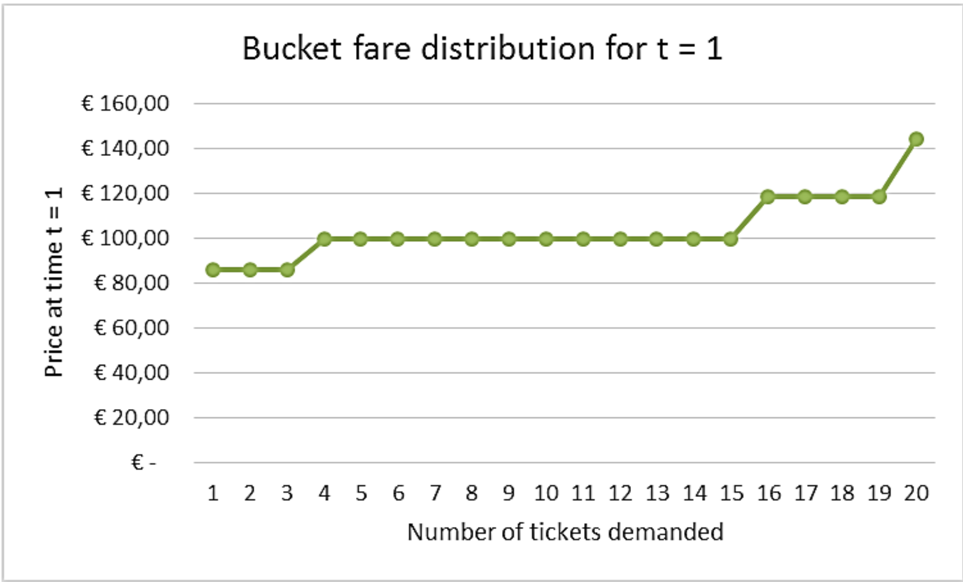

Figure 8: bucket fare distribution for t = 11 for the flight leaving in one week

On the other hand, the prices for the flight departing in one-month time are presented below in a 27 x 20 table. For graphical reasons, the table has been divided in half between column 10 and 11.

	x = 20	**x = 19**	**x = 18**	**x = 17**	**x = 16**	**x = 15**	**x = 14**	**x = 13**	**x = 12**	**x = 11**
t = 96	€ 60,19	€ 60,19	€ 71,81	€ 71,81	€ 71,80	€ 71,82	€ 71,76	€ 71,82	€ 85,98	€ 85,92
t = 95	€ 60,19	€ 60,19	€ 71,81	€ 71,81	€ 71,80	€ 71,82	€ 71,76	€ 71,82	€ 85,98	€ 85,92
t = 94	€ 60,19	€ 60,19	€ 71,81	€ 71,81	€ 71,80	€ 71,82	€ 71,76	€ 71,82	€ 85,98	€ 85,92
t = 93	€ 60,19	€ 60,19	€ 60,20	€ 60,18	€ 60,19	€ 71,87	€ 71,77	€ 71,73	€ 71,90	€ 71,78
t = 92	€ 60,19	€ 60,19	€ 60,20	€ 60,18	€ 60,19	€ 71,87	€ 71,77	€ 71,73	€ 71,90	€ 71,78
t = 91	€ 60,19	€ 60,19	€ 60,20	€ 60,18	€ 60,19	€ 71,87	€ 71,77	€ 71,73	€ 71,90	€ 71,78
t = 90	€ 60,19	€ 60,19	€ 60,20	€ 60,18	€ 60,19	€ 71,87	€ 71,77	€ 71,73	€ 71,90	€ 71,78
t = 89	€ 60,19	€ 71,81	€ 71,80	€ 71,80	€ 71,80	€ 71,86	€ 71,74	€ 85,96	€ 85,96	€ 85,88
t = 88	€ 60,19	€ 71,81	€ 71,80	€ 71,80	€ 71,80	€ 71,86	€ 71,74	€ 85,96	€ 85,96	€ 85,88
t = 87	€ 60,19	€ 71,81	€ 71,80	€ 71,80	€ 71,80	€ 71,86	€ 71,74	€ 85,96	€ 85,96	€ 85,88
t = 86	€ 60,19	€ 71,81	€ 71,80	€ 71,80	€ 71,80	€ 71,86	€ 71,74	€ 85,96	€ 85,96	€ 85,88
t = 85	€ 60,19	€ 60,19	€ 71,81	€ 71,81	€ 71,80	€ 71,82	€ 71,76	€ 71,82	€ 85,98	€ 85,92
t = 84	€ 60,19	€ 60,19	€ 71,81	€ 71,81	€ 71,80	€ 71,82	€ 71,76	€ 71,82	€ 85,98	€ 85,92
t = 83	€ 71,80	€ 71,80	€ 71,81	€ 71,79	€ 71,80	€ 71,84	€ 85,92	€ 85,96	€ 85,94	€ 85,94
t = 82	€ 71,80	€ 71,80	€ 71,81	€ 71,79	€ 71,80	€ 71,84	€ 85,92	€ 85,96	€ 85,94	€ 85,94
t = 81	€ 60,19	€ 71,81	€ 71,80	€ 71,80	€ 71,80	€ 71,86	€ 71,74	€ 85,96	€ 85,96	€ 85,88
t = 80	€ 60,19	€ 71,81	€ 71,80	€ 71,80	€ 71,80	€ 71,86	€ 71,74	€ 85,96	€ 85,96	€ 85,88
t = 79	€ 60,19	€ 71,81	€ 71,80	€ 71,80	€ 71,80	€ 71,86	€ 71,74	€ 85,96	€ 85,96	€ 85,88
t = 78	€ 60,19	€ 71,81	€ 71,80	€ 71,80	€ 71,80	€ 71,86	€ 71,74	€ 85,96	€ 85,96	€ 85,88
t = 77	€ 60,19	€ 71,81	€ 71,80	€ 71,80	€ 71,80	€ 71,86	€ 71,74	€ 85,96	€ 85,96	€ 85,88
t = 76	€ 60,19	€ 60,19	€ 71,81	€ 71,81	€ 71,80	€ 71,82	€ 71,76	€ 71,82	€ 85,98	€ 85,92
t = 75	€ 60,19	€ 60,19	€ 71,81	€ 71,81	€ 71,80	€ 71,82	€ 71,76	€ 71,82	€ 85,98	€ 85,92
t = 74	€ 60,19	€ 60,19	€ 71,81	€ 71,81	€ 71,80	€ 71,82	€ 71,76	€ 71,82	€ 85,98	€ 85,92
t = 73	€ 71,80	€ 71,80	€ 71,81	€ 71,79	€ 71,80	€ 71,84	€ 85,92	€ 85,96	€ 85,94	€ 85,94
t = 72	€ 71,80	€ 71,80	€ 71,81	€ 71,79	€ 71,80	€ 71,84	€ 85,92	€ 85,96	€ 85,94	€ 85,94
t = 71	€ 71,80	€ 71,80	€ 71,81	€ 71,79	€ 71,80	€ 86,00	€ 85,90	€ 85,90	€ 85,99	€ 86,01
t = 70	€ 71,80	€ 71,80	€ 71,81	€ 71,79	€ 71,80	€ 86,00	€ 85,90	€ 85,90	€ 85,99	€ 86,01

Table 7: Actual prices for 20 seats collected for the Easyjet Rome – Amsterdam flight departing in one month.

	x = 10	x = 9	x = 8	x = 7	x = 6	x = 5	x = 4	x = 3	x = 2	x = 1
t = 96	€ 85,97	€ 85,93	€ 85,87	€ 85,95	€ 99,68	€ 99,42	€ 99,75	€ 99,69	€ 99,43	€ 99,61
t = 95	€ 85,97	€ 85,93	€ 85,87	€ 85,95	€ 99,68	€ 99,42	€ 99,75	€ 99,69	€ 99,43	€ 99,61
t = 94	€ 85,97	€ 85,93	€ 85,87	€ 85,95	€ 99,68	€ 99,42	€ 99,75	€ 99,69	€ 99,43	€ 99,61
t = 93	€ 71,74	€ 85,98	€ 85,93	€ 85,89	€ 86,06	€ 85,84	€ 86,08	€ 99,58	€ 99,51	€ 99,59
t = 92	€ 71,74	€ 85,98	€ 85,93	€ 85,89	€ 86,06	€ 85,84	€ 86,08	€ 99,58	€ 99,51	€ 99,59
t = 91	€ 71,74	€ 85,98	€ 85,93	€ 85,89	€ 86,06	€ 85,84	€ 86,08	€ 99,58	€ 99,51	€ 99,59
t = 90	€ 71,74	€ 85,98	€ 85,93	€ 85,89	€ 86,06	€ 85,84	€ 86,08	€ 99,58	€ 99,51	€ 99,59
t = 89	€ 86,01	€ 85,87	€ 85,93	€ 99,69	€ 99,65	€ 99,49	€ 99,50	€ 99,84	€ 99,53	€ 118,09
t = 88	€ 86,01	€ 85,87	€ 85,93	€ 99,69	€ 99,65	€ 99,49	€ 99,50	€ 99,84	€ 99,53	€ 118,09
t = 87	€ 86,01	€ 85,87	€ 85,93	€ 99,69	€ 99,65	€ 99,49	€ 99,50	€ 99,84	€ 99,53	€ 118,09
t = 86	€ 86,01	€ 85,87	€ 85,93	€ 99,69	€ 99,65	€ 99,49	€ 99,50	€ 99,84	€ 99,53	€ 118,09
t = 85	€ 85,97	€ 85,93	€ 85,87	€ 85,95	€ 99,68	€ 99,42	€ 99,75	€ 99,69	€ 99,43	€ 99,61
t = 84	€ 85,97	€ 85,93	€ 85,87	€ 85,95	€ 99,68	€ 99,42	€ 99,75	€ 99,69	€ 99,43	€ 99,61
t = 83	€ 85,95	€ 85,93	€ 99,52	€ 99,64	€ 99,61	€ 99,55	€ 99,58	€ 99,64	€ 118,24	€ 118,14
t = 82	€ 85,95	€ 85,93	€ 99,52	€ 99,64	€ 99,61	€ 99,55	€ 99,58	€ 99,64	€ 118,24	€ 118,14
t = 81	€ 86,01	€ 85,87	€ 85,93	€ 99,69	€ 99,65	€ 99,49	€ 99,50	€ 99,84	€ 99,53	€ 118,09
t = 80	€ 86,01	€ 85,87	€ 85,93	€ 99,69	€ 99,65	€ 99,49	€ 99,50	€ 99,84	€ 99,53	€ 118,09
t = 79	€ 86,01	€ 85,87	€ 85,93	€ 99,69	€ 99,65	€ 99,49	€ 99,50	€ 99,84	€ 99,53	€ 118,09
t = 78	€ 86,01	€ 85,87	€ 85,93	€ 99,69	€ 99,65	€ 99,49	€ 99,50	€ 99,84	€ 99,53	€ 118,09
t = 77	€ 86,01	€ 85,87	€ 85,93	€ 99,69	€ 99,65	€ 99,49	€ 99,50	€ 99,84	€ 99,53	€ 118,09
t = 76	€ 85,97	€ 85,93	€ 85,87	€ 85,95	€ 99,68	€ 99,42	€ 99,75	€ 99,69	€ 99,43	€ 99,61
t = 75	€ 85,97	€ 85,93	€ 85,87	€ 85,95	€ 99,68	€ 99,42	€ 99,75	€ 99,69	€ 99,43	€ 99,61
t = 74	€ 85,97	€ 85,93	€ 85,87	€ 85,95	€ 99,68	€ 99,42	€ 99,75	€ 99,69	€ 99,43	€ 99,61
t = 73	€ 85,95	€ 85,93	€ 99,52	€ 99,64	€ 99,61	€ 99,55	€ 99,58	€ 99,64	€ 118,24	€ 118,14
t = 72	€ 85,95	€ 85,93	€ 99,52	€ 99,64	€ 99,61	€ 99,55	€ 99,58	€ 99,64	€ 118,24	€ 118,14
t = 71	€ 85,83	€ 99,69	€ 99,50	€ 99,54	€ 99,64	€ 99,48	€ 99,78	€ 118,20	€ 118,35	€ 118,19
t = 70	€ 85,83	€ 99,69	€ 99,50	€ 99,54	€ 99,64	€ 99,48	€ 99,78	€ 118,20	€ 118,35	€ 118,19

Table 7 bis: Actual prices for 20 seats collected for the Easyjet Rome – Amsterdam flight departing in one month.

Below, the main descriptive statistics for the above 27 x 20 matrix are shown.

Mean		€ 84,76
Median		€ 85,93
St. Deviation		14,37
Variance		206,49
Min		€ 60,18
Max		€ 118,35
Range		€ 58,17
1st Quartile		€ 71,80
3rd Quartile		€ 99,50

Table 8: Descriptive Statistics for the Easyjet Roma-Amsterdam flight departing in one month

As for the fare distribution for the flight departing in one month, the statistical distribution appears to be rather symmetrical with mean and median almost coinciding. In addition, the standard deviation is not particularly significant, at least compared to the previous series of observation, indicating less fluctuation in the values. This is also highlighted by the lower range, the difference between the fourth and the zero quartile. Nevertheless, surprisingly, the minimum value in this series of observation is higher than the one observed before and this may be explained by the fact that the company has not undergone an attractive pricing policy by further lowering the lowest buckets. On the other hand, as expected, the maximum value is lower than the one for the other flight as business travellers and strategic consumers have not been targeted yet.

Below graphs for t = 96, t = 88, t = 79 and t =70 can be found. On the x-axis, I consider the number of tickets for which the query has been made. On the other hand, on the y-axis, I plot the price as a function of x.

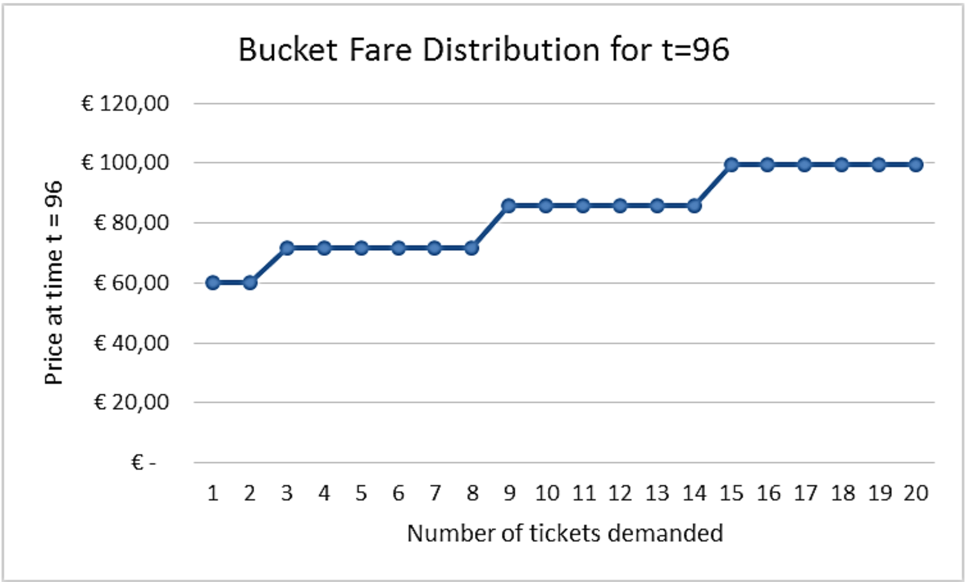

Figure 9: bucket fare distribution for t = 96 for the flight leaving in one month

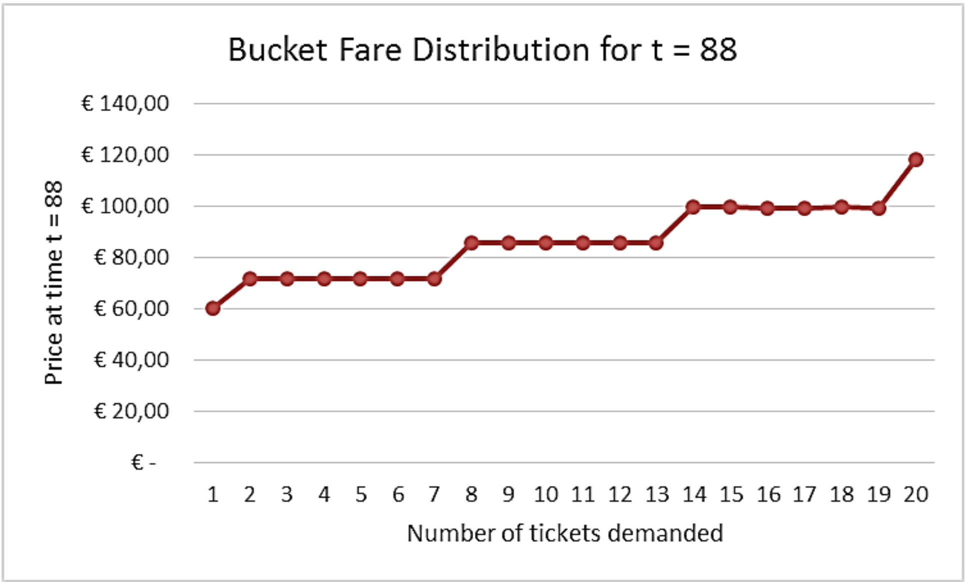

Figure 10: bucket fare distribution for t = 88 for the flight leaving in one month

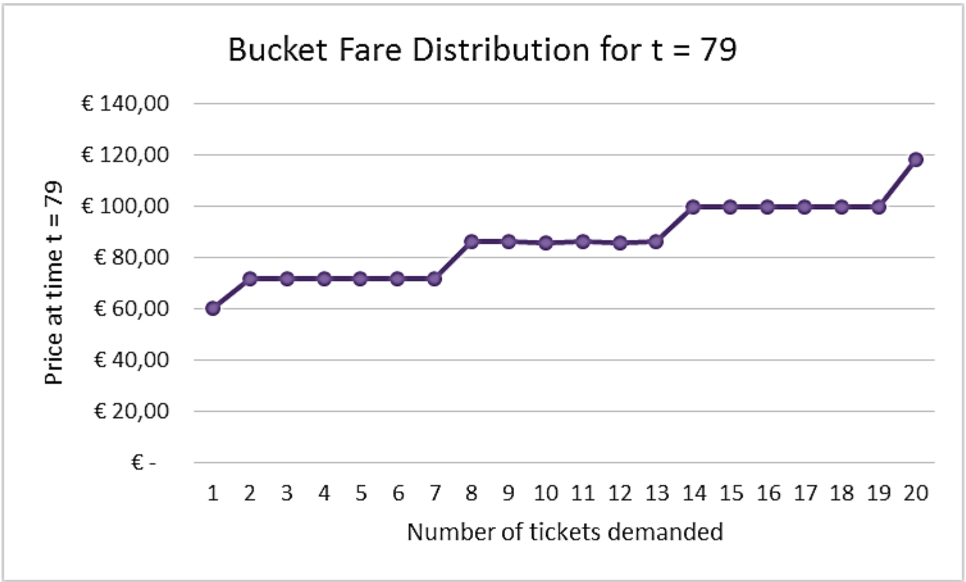

Figure 11: bucket fare distribution for t = 79 for the flight leaving in one month

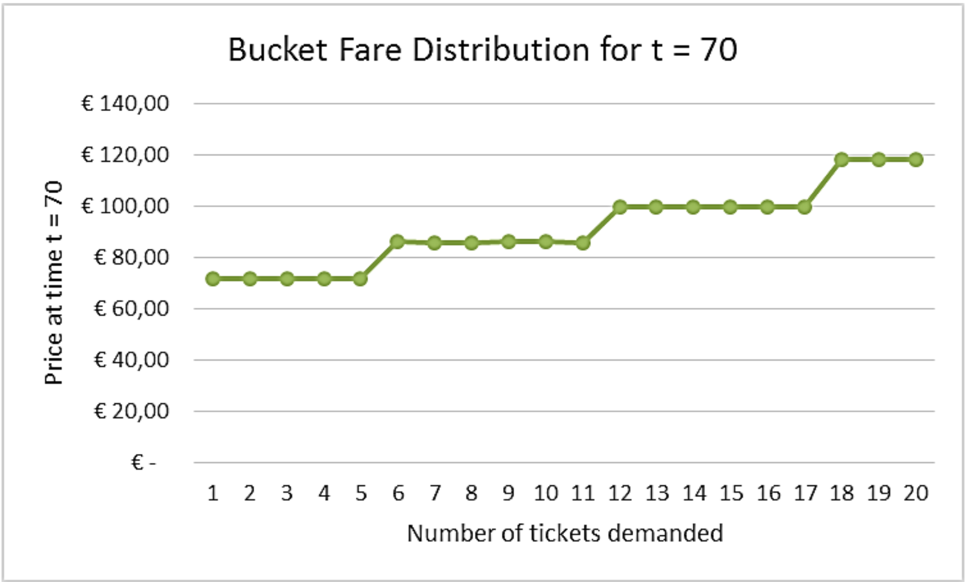

Figure 12: bucket fare distribution for t = 70 for the flight leaving in one month

6. CONCLUSION

This work provides a set of important and partially innovative contributions to the existing literature on airline pricing by delineating a possible approach to make the airline reservation system and the functioning of dynamic pricing algorithms more transparent for consumers.

First, I have presented a theoretical model that aims at providing customers adequate tools to understand how carriers actually implement their dynamic pricing algorithms, thus allowing, under certain assumptions, to derive a prediction of how the pricing distribution is likely to evolve over time. Indeed, as in Alderighi, Gaggero and Piga (*The hidden side of dynamic pricing in airline markets*, 2016), I expect the carrier to be able to modify the distribution as the date of departure approaches, thus allowing to derive a prediction of how the distribution is likely to evolve over time. Similar to the various models of revenue management surveyed in McAfee and te Velde (2007), my model has predicted an intrinsic declining value of seats as the departure date approaches. Nonetheless, to avoid strategic behaviour from consumers who tend to postpone the purchase of their tickets as well as to exploit the higher willingness-to-pay of business travellers, the price observed by prospective buyers is expected to increase as the date of departure nears.

Second, I have claimed that, in order to understand how flight prices evolve over time, focusing on the fare for the first seat on sale is not a valid way to conduct a test on the fare intertemporal distribution for several reasons. For instance, consistent with the airline and revenue management literature, my model has predicted the fare distribution to be monotonically increasing. Therefore, if each seat in a flight is assigned a different fare, then tracking the fare time path of the first seat on sale implies often tracking the fare of different seats over time. Conversely, when trying to predict the fare distribution of a given flight, it is useful to consider the different buckets of seats as a whole, where each bucket is defined by an increasing price tag and a variable size, as empirically shown in 5.3.

By blending theory and practice in different fields, I have replicated the functioning of the dynamic pricing algorithm used by Alderighi, Gaggero and Piga (*The hidden side of dynamic pricing in airline markets*, 2016). The prices resulting from this algorithm's simulation have been compared to actual prices charged by Easyjet for the route Rome-Amsterdam, as a proof of concept of its functioning. In particular, it has been shown that my algorithm is quite reliable in predicting the actual fares. However, it can be argued that in order to gain a more precise price assessment a variable probability of arrival can be considered and a seasonality parametric factor introduced. Moreover, another arguable limitation of the model is that, for the sake of simplicity, the consumer was assumed to observe one and only one price p at each time t. Nevertheless, as already mentioned in 4.1, this is not the case as far as online queries on Easyjet's website are concerned. Last, for a more in-depth analysis on the topic, a parametrization of my algorithm for the actual case, that is, for the 132 seats available on the plane, as well as for the 6 months preceding the take-off, should be considered.

Furthermore, in 5.3 I have shown the bucket fare distribution under two real world scenarios. The main limitation of this analysis is that the variable x is fixed to 20: a non-negligible constraint for the reliability of the study. However, future scholars should bear in mind that, due to the way the Easyjet website works, it is not possible to buy more than 40 tickets for a given query and that the data retrieval operation in a 6-month window can be rather daunting, as highlighted in Section 4.

My analysis provides important insights on how understanding dynamic pricing may be helpful for customers dealing with such firms as hotels, cruise ships, car rentals, which set their prices facing conditions similar to those of airlines. With regard to the airline industry, the present work represents an initial step towards providing customers a view of how revenue management techniques are implemented by Low Cost Carriers (LCCs), such as Easyjet. Nevertheless, future research needs to investigate how my findings can be adapted to the more complex revenue management systems adopted by traditional carriers as well as by other major firms operating in the transportation industry (e.g. Trenitalia, Costa Crociere).

7. REFERENCES

Abrate, G., Fraquelli, G., & Viglia, G. (2012). Dynamic pricing strategies: Evidence from european hotels. *31, 1*, 160-168. International Journal of Hospitality Management.

Alderighi, M., Gaggero, A., & Piga, C. A. (2016). The hidden side of dynamic pricing in airline markets. MPRA Paper.

Alderighi, M., Nicolini, M., & Piga, C. A. (2016). Targeting leisure and business passengers with unsegmented pricing. *50*, 502-515. Tourism Management.

Borenstein, S., & Rose, N. L. (1994). Competition and Price Dispersion in the US Airline Industry. *102*, 653-683. Journal of Political Economy.

Dana, J. D. (1999). Equilibrium Price Dispersion under demand uncertainty: the roles of costly capacity and market structure. *30, 4*, 632-660. Rand Journal of Economics.

Deneckere, R., & Peck, J. (2012). Dynamic Competition with Random Demand and Costless Search: a Theory of Price Posting. *80, 3*, 1185-1247. Econometrica.

Gaggero, A. A., & Piga, C. A. (2010). Airline Competition in the British Isles. *46*, 270-279. Transportation Research Part E: Logistics and Transportation Review.

Gaggero, A. A., & Piga, C. A. (2011). Airline Market Power and Intertemporal Price Dispersion. *59, 4*, 552-577. Journal of Industrial Economics.

Lazarev, J. (2013). The Welfare Effects of Intertemporal Price Discrimination: An Empirical Analysis of Airline Pricing in U.S. Monopoly Markets. .

Li, J., Granados, N., & Netessine, S. (2014). Are Consumers Strategic? Structural Estimation from the Air-Travel Industry. *60*, 2114-2137. Management Science.

McAfee, R. P., & te Velde, V. (2007). Dynamic pricing in the airline industry. In T. J. Hendershott, *Handbook of Economics and Information Systems* (Vol. 1). New York(NY): Elsevier Science.

Moller, M., & Watanabe, M. (2010). Advance Purchase Discounts and Clearance Sales. *120*, 1125-1148. Economic Journal.

Puller, S. L., Sengupta, A., & Wiggins, S. (2009, December). Testing theories of Scarcity pricing and Price Dispersion in the Airline Industry. NBER Working Paper Series.

Sheppard, L. (2018). *A machine-learning approach to inventory-constrained dynamic pricing*. Retrieved from MIT News: http://news.mit.edu/2018/thompson-sampling-machine-learning-inventory-constrained-dynamic-pricing-0315

Smyth, B. (2015, May). *Dynamic Pricing And Price Discrimination: What's The Difference?* Retrieved from Wiser: https://blog.wiser.com/dynamic-pricing-and-price-discrimination-whats-the-difference/

Sweeting, A. (2012). Dynamic Pricing Behavior in Perishable Goods Markets. *Evidence from Secondary Markets for Major League, 120, 6*, 1133-1172. Journal of Political Economy.

Talluri, K. T., & van Ryzin, G. J. (2004). The Theory and Practice of Revenue Management. New York, United States: Springer.