Maximizing Airline Profit

Rebecca Derham, Madison Guzman, Jaimie Lee, Oluwatosin Olajubu

December 18, 2023

Abstract

Dynamic pricing is a powerful strategy that adapts the price of a product or service to changing supply and demand conditions. In this paper, we present a dynamic pricing model for airline tickets that takes into account the passenger type and the date at which they purchase a ticket of a certain class. Our model is a nonlinear program that maximizes the revenue of a single flight. We use demand curves that capture how different customers value different ticket classes and how their willingness to pay changes over time. Using parameters that aim to simulate the demand of a given flight, our model produces results that are consistent with industry averages. The model delivers an optimal pricing schedule that matches our expectations of how an airline's revenue is distributed across the different ticket classes. Therefore, our model can help airlines improve their pricing strategies as well as help customers make more informed decisions on when to purchase their airfare. We also discuss the limitations of our model and propose avenues for future research.

Contents

1 Introduction

1.1 Problem Statement

Historically, the airline industry used to be much simpler to understand. Each seat on a given flight path was always priced using a flat rate. However, in the 1980s, American Airlines became the first airline to adopt a new ticket pricing strategy known as dynamic pricing [1]. In this system, airlines would offer cheaper rates to passengers who book earlier, offer seat reservations at a higher price, and incorporate the now controversial practice of overbooking. This was a revolutionary process at the time. To accomplish this, American Airlines developed a system where flight prices would change in real time based on timely data such as customer booking patterns, competitor prices, and even weather and popular events that may impact demand. Shortly after, the entire airline industry adopted this dynamic pricing model. This pricing model is now industry-standard and has become more complex than ever. Given today's technology, airlines have been able to gather large amounts of data to help fine-tune aspects of their model. This improved information allows them to better adjust prices in response to changes in consumer behavior or general supply and demand imbalances.

As a result, consumers are always looking for the best price. Nowadays, there exist several websites, such as Kayak or Google Flights, that help guide customers on determining the right time to buy a certain flight. They attempt to predict when airlines will either increase or decrease their prices, and give customers a suggestion on whether to "buy now" or "buy later". We want to understand how these websites are able to predict these price changes. To do so, we will investigate how airlines determine their pricing schedule based on a variety of factors. We will attempt to build a similar dynamic pricing model that follows the characteristics we described above and that allows us to determine a specific pricing strategy to maximize an airline's profit.

1.2 Summary of Our Approach

To build this model, we first need to determine its structure. The objective function we want to maximize is the airline's profit for a given flight. In our case, this profit will be the sum of the total tickets sold at a given point in time multiplied by the price of the ticket at that time. However, our problem is far more complicated than that. Once we determine the different time periods at which we will change prices, we need to decide the rate at which we change prices. We also need to decide what other factors we will consider in pricing the tickets; for example, it's typically the case that first class tickets are more expensive than economy tickets. In addition, airlines might also require that a certain number of seats be left empty so they can capitalize on last-minute flyers who are willing to pay more for their tickets. On top of that, a more complex concept airlines must determine is the passenger types who are interested in a given flight. Certain passenger types might have different buying behavior at different times before the departure date. For example, a person flying for leisure purposes might be more or less sensitive to price changes at different time periods compared to a passenger flying for business purposes. We attempt to include as many of these ideas as possible into our model.

Once we have our model formulation, complete with the objective function and constraints, we will use the Python's SciPy library to find a solution. This involves listing the parameters of our model, each constraint, and the objective function. One function call allows us to combine these and output a solution, which will tell us what the price of a certain ticket class should be at a given time.

2 Background Research

Before we begin, it is worth learning more about the current state of the airline ticketing industry. Airlines typically start accepting bookings for flights around 11 months before the flight departure date [2]. Prices decrease about four months before the departure date. Three to four weeks before the departure date, the prices start to increase. This comes as airlines anticipate more business travelers to book during this time since they (or their companies) are willing to pay more for air travel on a short notice. This provides us with useful information to help us determine at which time points we should schedule a price change. To generalize from the information we found above, we will use the following times: 1 year, 6 months, 1 month, and 1 week away from the departure date. However, it is important to note that when airlines are making the decision to change prices, they do so at a greater frequency. That is, their model allows for real-time price changes in response to a variety of factors. However, for simplicity, our model will only include these four critical time periods.

The timing system above provides us with a general outline of the model we will build. However, it adds only one layer of complexity to our model. More research informed us of what else would be needed in our model. One such element is the the now common practice of overbooking [3]. Airlines will purposely overbook their flight to ensure that there are no empty seats. Airlines bet that a handful of passengers either cancel or don't show up to their flight. This is a gamble they are willing to make as the lost revenue associated with each seat left empty by a no-show is more than just the cost of that seat; it includes all of the lost revenue associated with potential passengers who were willing to buy that seat but were ultimately turned away and whose business went to another airline. Our model will attempt to simulate this overbooking by having an upper bound on the number of seats that is some percentage above the capacity of the plane.

Additionally, we also wanted to include another common practice of airlines: "protecting" seats on each flight [3]. Airlines protect (or reserve) a group of seats on every flight for "full fare" flyers. This is so that last-minute flyers willing to pay last-minute prices can buy these more expensive seats and increase the revenue for that flight. However, airlines must make the difficult decision of choosing how many seats to protect and for how long. In order to do this, airlines do not simply reserve a certain number of seats across the entire capacity of their plane. Instead, they section their plane by the type of ticket fare (first class, business, economy, etc.). To determine how many seats of each ticket type they need to protect, an airline needs to find a way to accurately measure demand for each of these ticket types. To do so, airlines must do their research on the variety of different passenger types. Each passenger type might respond differently to price changes at different time points.

To model these different types of passengers, we conducted research on passenger demographics. A 2016 survey found that more than 31% of people traveled for business purposes while 48% traveled for leisure [4]. The remaining 21% traveled for personal non-leisure purposes. This information is useful for us as we will use it to build passenger profiles. Certain passenger types might be more likely to buy tickets at a certain time period and might be more or less sensitive to higher prices. This could also translate to certain passenger types being more or less likely to buy fares in certain classes. For example, we can think about the case of business travelers. These business travelers might be more inclined to fly in first or business class because their company will finance their ticket. Additionally, because business trips might arise on short notice, flyers may be more willing to pay higher prices to claim their seat. On the other hand, passengers flying for leisure purposes, such as for a vacation, may book their vacation far in advance. If they're booking far in advance, they might want to take advantage of the cheaper seats in economy class at that time. This range of consumer behavior shows how having this segmentation is necessary in our model. We will use this partitioning of passenger types and our own intuition to develop probability distributions on how likely groups of passengers are to buy tickets at a given price.

More research also informed us of what was not needed in our model. Researchers at the National Bureau of Economic Research built a theoretical model to simulate a dynamic pricing system for airfare tickets [5, 6]. They found that dynamic pricing systems that factor in competitor behavior could result in airlines selling too many tickets too soon. As competition is added, the problem becomes infinitely harder as everybody starts reacting to each other. If Airline A drops its price, Airline B responds by also dropping its price; over time, the two airlines will bid each other down to the point where they sell too many tickets at these very low prices. This hurts airlines that rely on the high prices last-minute customers are willing to pay, as well as customers who need that next-day flight. However, systems that ignore competitors and instead rely on internally determined pricing rules can avoid that trap, benefiting consumers and companies alike. Researchers found that this approach results in 4-5% revenue increase and 3% consumer surplus increase when compared to their theoretical system that factored in competitor's prices. Therefore, we choose to have our dynamic pricing model not include any influence from competitors.

3 Formulation

3.1 Objective Function

With this background information, we now look towards formulating our model. We will use the following time points at which we will designate a price change: 1 year, 6 months, 1 month, and 1 week from the departure date. Additionally, we will use the following ticket classifications: first class, business class, and economy class. We also need to denote the different customer types: business, leisure, and personal travelers. We introduce these ticket types and passenger types to provide a more realistic model of ticket demand. It is known that different fares are priced differently. From the research conducted, we know that different ticket classes may attract different passenger types, and so airlines respond by reserving a number of seats in each class to capitalize on consumer behavior.

In addition to these classifications defined above, we need to define two parameters. In order to model consumer behavior, we will first define demand functions for each customer type at each time point leading up to the departure date. These demand functions will be a function of a price, and it will tell us the likelihood that a specific type of customer is willing to pay a given price for a certain ticket type at a time point. We will also define market size parameters, which will indicate the total number of each customer type who are interested in purchasing a ticket at some time point. Further details on how these parameters were defined for our model are outlined in our "Parameters" section.

With our model, we aim to solve for a price schedule that defines how ticket prices of each class will change at different points in time leading up to the flight's departure date. Therefore, our objective function will maximize the expected profit given the price schedule and the expected number of tickets sold at a given time. The expected number of tickets sold will be calculated using the demand function and market size parameters. Below is a derivation of our model's objective function.

$$
\begin{aligned}
\max \text{Profit} &= \max \sum_{\text{Time } t} \mathbb{E}[\text{profit at time } t] \\
&= \max \sum_{\text{Time } t} \sum_{\text{Ticket } i} \text{Price} \times \ \mathbb{E}[\#\text{tickets of type } i \text{ sold at time } t \mid \text{Price}] \\
&= \max \sum_{\text{Time } t} \sum_{\text{Ticket } i} \text{Price} \times \left(\sum_{\text{Customer } c} (\#\text{customers of type } c \text{ at time } t) \ \times \ (\%\text{willing to pay at least Price}) \right)
\end{aligned}
$$

For the subsequent sections, we use the following abbreviations for each component defined in our model.

time	$t \in [1, 2, 3, 4]$
ticket types	$i \in [f, b, e]$
customer types	$c \in [B, P, L]$
demand function of customer type for ticket type at time	$f_{c,i,t}$
market size of customer type at time	$n_{c,t}$
price of ticket type at time	$P_{i,t}$

3.2 Constraints

With the basic set-up of our model completed, we now introduce the three categories of constraints we incorporated.

3.2.1 Price Ordering

To start, we want the prices of the ticket classes at each time point to follow an intuitive ordering. That is, at each time point, first class tickets should be more expensive than business class tickets which should be more expensive than economy class tickets.

$$P_{f,t} \geq P_{b,t} \geq P_{e,t} \qquad \forall t \in [1, 2, 3, 4]$$

3.2.2 Capacity

We also include a constraint that checks how full a flight is once all the seats have been sold. Summing over all four time periods, we must have that the expected number of seats sold across each class is less than or equal to an upper bound of seats for each class.

This upper bound may or may not allow for overbooking, and we will allow the option to include it once we establish our parameters. It will therefore be equal to the total number of seats available in that fare class with or without an additional percentage above that bound to allow for overbooking.

$$\sum_{t=1}^{4} \sum_{c \in [B,P,L]} n_{c,t} \cdot f_{c,i,t}(P_{i,t}) \leq L_i \qquad \forall i \in [f, b, e]$$

3.2.3 Target Booking

We also want to establish a time-specific booking constraint that models how airlines reserve a certain number of seats in each fare class. To do so, we require that the expected number of seats of each class sold at or before a given time period must be between a minimum and maximum bound. This allows us to set benchmark booking goals at each time period for each class so that airlines can reserve seats at different times to match consumer behaviors.

$$m_{i,t} \leq \sum_{t'=1}^{t} \sum_{c \in [B,P,L]} n_{c,t'} \cdot f_{c,i,t'}(P_{i,t'}) \leq M_{i,t} \qquad \forall t \in [1, 2, 3, 4], \forall i \in [f, b, e]$$

3.3 Parameters

We now expand upon how we defined the following parameters when running our model as well as any reasoning.

3.3.1 Demand Functions

As shown in the derivation of our objective function, we chose to use the expected number of tickets sold so that we can introduce different customer types. Each passenger type will have a different probability distribution for each fare type in order to model their behavior for a given price point. That is, these demand functions will tell us how likely a customer of a given type is willing to buy a seat of a certain class at a given time period for a given price. These functions are illustrated in Figure 1 on the next page.

We now explain our intuition for choosing these functions in particular, beginning with those for customers traveling for leisure. First, the functions change over time in a similar manner for all seat types; they begin somewhat high a year in advance, as if someone is booking a leisure trip a year in advance they are likely pretty sure they want to take the trip. The prices then decrease 6 months and 1 month in advance, before increasing a lot 1 week in advance; when booking a trip a week beforehand, there is a lot less flexibility which results in people being willing to pay higher prices. Finally, while people in this category are willing to pay slightly higher prices for better seats, our demand functions do not reflect too drastic of a preference; the majority of people traveling for leisure will not want a better seat enough to be willing to pay a premium.

Our demand functions for customers traveling for personal but non-leisure reasons generally follow similar trends. However, planning non-leisure travel generally involves more constraints than leisure travel does, such as needing to travel on specific dates. As a result, people will likely be willing to pay more for a given ticket, and also have a stronger cutoff for the maximum price they are willing to pay (resulting in a

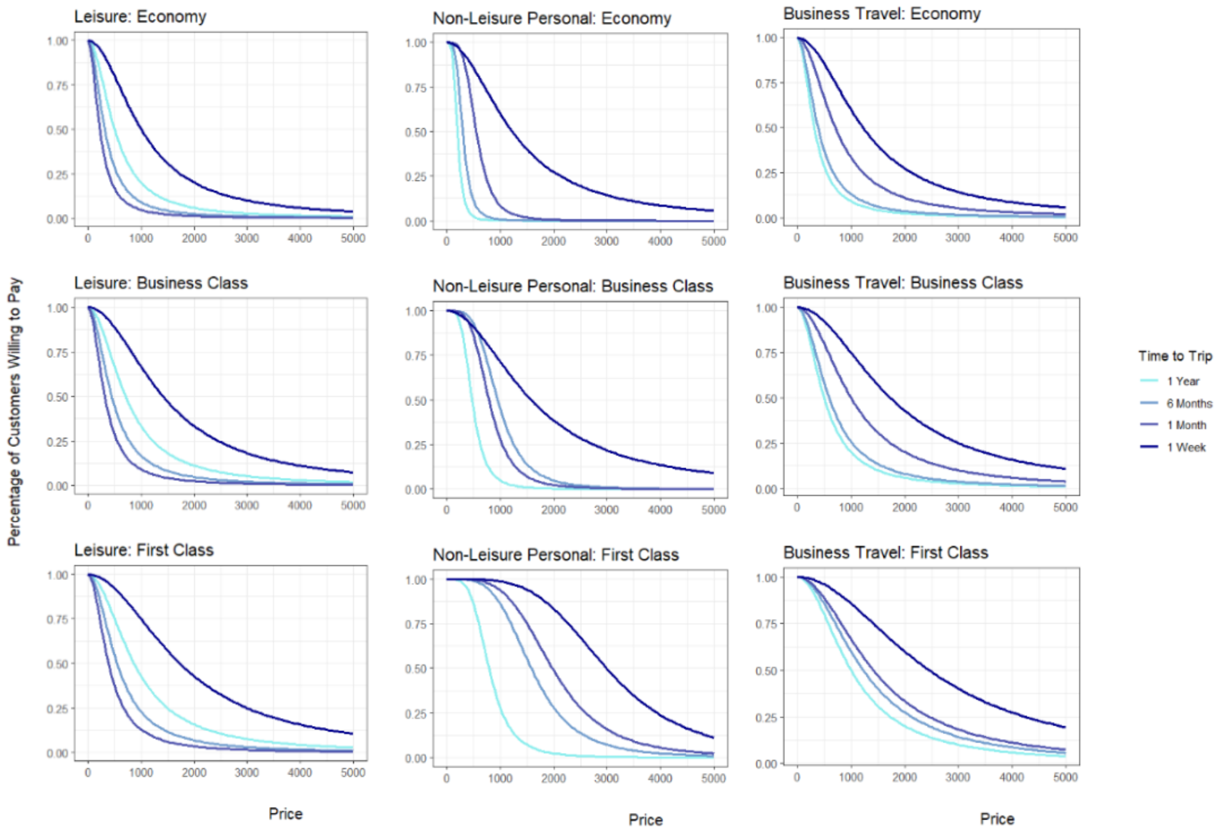

Figure 1: Demand Functions

steeper graph than the previous category). In addition, these customers are less likely to be traveling with their children which means they are more willing to pay a higher price for a better ticket; this is reflected in a larger difference in demand curves between different seat types. Finally, as non-leisure travel includes personal emergencies, which would be booked within a week of the flight, people are willing to pay much higher prices closer to the flight, with less of a sharp cutoff, which is reflected in a smoother, flatter curve.

Finally, we discuss the demand curves for business-related travel. There are three main differences between this category and the others. First, these trips are necessary and as a result the curves are flatter; more people will be willing and able to pay a very high price for a necessary business trip than for leisure or personal travel. Second, business trips are very rarely booked more than 1 month in advance of the trip, with nearly 50% of business trips booked within a week of the flight [7]; this results in lower prices 1 year or 6 months in advance. Finally, business travelers have the strongest preference for better tickets, and so there is the greatest difference in the demand curves across seat types.

3.3.2 Market Sizes

Based upon our research, we estimated the ratio of each type of customer looking to purchase flight tickets at each time point. Using these estimated ratios, we defined the following market size parameters.

	1 year	6 months	1 month	1 week
Business	3	4	4	7
Personal	7	10	20	30
Leisure	25	40	40	10

3.3.3 Upper Bound for Capacity Constraints

As for our capacity constraints, we decided to allow for 0%, 2%, and 4% overbooking in first, business and economy class seats, respectively. To define the upper bounds for our constraints, we used the A320 plane, whose seat count is listed below.

	Number of Seats	Maximum Seats Booked, L_i
First Class	12	12
Business Class	42	43
Economy Class	96	100

3.3.4 Minimum and Maximum Bounds for Target Booking Constraints

Approximating based on consumer behavior and market sizes, we defined the following minimum and maximum bounds on how many cumulative tickets our model aims to sell by each time point.

	1 year	6 months	1 month	1 week
First Class	[0,3]	[2,6]	[5,9]	[8,12]
Business Class	[6,18]	[15,33]	[25,38]	[38,43]
Economy Class	[20,46]	[36,78]	[64,94]	[90,100]

3.4 Final Formulation

Now, with each part of our model defined, we have the following LP formulation:

$$\max \sum_{t=1}^{4} \left(\sum_{i \in [f,b,e]} P_{i,t} \cdot \left(\sum_{c \in [B,P,L]} n_{c,t} \cdot f_{c,i,t}(P_{i,t} \right) \right)$$

such that

$$
\begin{aligned}
P_{f,t} \geq P_{b,t} \geq P_{e,t} & \qquad \forall t \in [1,2,3,4] \\
\sum_{t=1}^{4} \sum_{c \in [B,P,L]} n_{c,t} \cdot f_{c,i,t}(P_{i,t}) \leq L_i & \qquad \forall i \in [f,b,e] \\
m_{i,t} \leq \sum_{t'=1}^{t} \sum_{c \in [B,P,L]} n_{c,t'} \cdot f_{c,i,t'}(P_{i,t'}) \leq M_{i,t} & \qquad \forall t \in [1,2,3,4], \forall i \in [f,b,e]
\end{aligned}
$$

4 Solution

Our LP formulation was implemented using the `scipy.optimize` package in Python. Upon running, the optimization terminated successfully. Our results are shown in Figures 2 and 3.

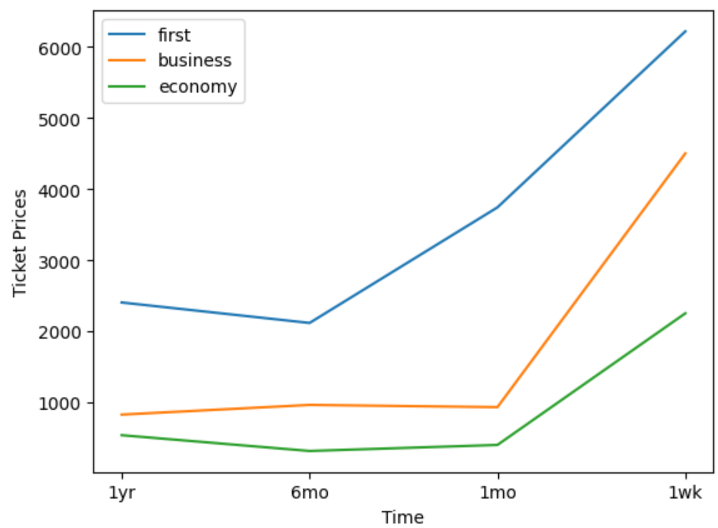

Figure 2: Pricing Schedule

	1 year	6 months	1 month	1 week
First Class	\$2,400	\$2,100	\$3,750	\$6,200
Business Class	\$830	\$970	\$930	\$4,500
Economy Class	\$540	\$320	\$400	\$2,200

Figure 3: Ticket Prices by Class and Time

We make a few observations regarding the resulting price schedule. For the first class tickets, there is a price drop from the 1 year-time-point to the 6 months-time-point. Starting from the 6 month-time-point, there is a relatively steady increase in prices leading up to the flight's departure date. For the business class tickets, there is a slight increase then decrease between the 1 year-time-point and the 1 month-time-point. From this point, relative to the increase in first class prices, we see a steeper increase in the business class ticket price. Lastly, for the economy class tickets, there is a slight decrease then increase between the 1 year-time-point and the 1 month-time-point. From the 1 month-time-point and onward, there is a steeper increase in prices, however, the slope of this increase is not as steep as that of the first class and the business class. In generalizing these results, ticket prices for each ticket class are the lowest in between the 6 months-time-point and 1 months-time-point before the departure date. This observation is consistent with what can be seen in the current market for purchasing flight tickets.

The total profit made under this pricing schedule given the assumptions we made regarding market sizes and consumer behavior is $140,961.

5 Limitations & Next Steps

Our research aimed to explore how flight-finding websites, such as Kayak, utilize airline data to forecast how airlines set the prices of flight tickets. To this end, we developed a dynamic pricing model that would generate an optimal price schedule for flight tickets that maximizes an airline's profit.

Before we made the decision to use our own parameters to simulate the demand over time for a given flight, we attempted to use real market data. However, once we found a reliable dataset, we realized that its size exceeded 28GB. In order to work with this data, we would need a computer with significantly larger processing capabilities than our own laptops. Rather than go down this route, we decided to instead use our own parameters to simulate this real-world data. To do so, we conducted research on existing airline strategies and the overall distribution of passengers on a given flight. This means that our model is based on the literature review of the airline industry and the characteristics we learned about customer behavior. We also note that airlines collect real-time data and therefore have the ability to adjust prices more frequently based on the latest changes in demand as well as very fine-tuned consumer behavior that has been complied over years. However, because we don't have access to data of that scale, we decided to use demand functions that simulate customer behavior according to our literature review and assumptions. In future work, we'd like to use these larger data sets and determine whether our model works just as well using real-world data versus our own simulated parameters.

In addition, our model makes the simplifying assumption that cost is constant. That is, when we aim to maximize an airline's profit, our model is really only maximizing their revenue. However, when airlines create their pricing strategies, they must include any costs incurred in chartering any given flight. One of the factors that influences the fuel costs for an airline is the weight of the aircraft and its contents. The heavier an aircraft is, the more fuel it requires to fly, which increases the operational costs and carbon emissions. OpenAirlines reports that a typical cost of weight is 3.5 % per hour of flight, which means that every kilogram or pound saved contributes directly to reducing fuel consumption [8]. National Geographic estimates that every laptop, pillow, or magazine that passengers bring along adds to the hidden costs of flying [9]. Therefore, airlines are constantly looking for ways to reduce weight by changing the design and components of their aircraft [10].

Another limitation of our model is that it only determines prices for one flight and does not take into account the distance, the departure location, or the destination of the flight. These are all significant factors that affect the ticket prices and the overall flight costs. Especially during peak seasons such as spring, summer, or winter breaks, certain destinations are in greater demand and would consequently be priced at a higher rate. Although we could not test our model with real data, we believe that the price schedule we obtained is consistent with industry averages. However, we recognize that these factors are indeed important in the pricing of a flight. In the future, with more data, we'd be able to use better judgement in setting the parameters for different flight paths and at different times of the year.

Lastly, in later work, we would like to incorporate competition into our pricing model. According to a paper by researchers at Yale, including competition in a pricing model leads to a lower overall profit [6]. This is because competitors respond to each other and lower their bids, which causes seats to be sold too fast and at reduced prices. We want to apply competition to our model and compare our outcomes with the findings of their study. If we also find that adding competition leads to a reduced profit, we'd have confidence that our model is producing similar results and have an extra sense of affirmation that competition should not be added into our model.

Overall, these avenues for future work showcase the sheer complexity of airline pricing models. Accurately modeling consumer behavior is hard as it is. Airlines try their best to adjust their prices to take advantage of different behaviors across different passenger types as well as changing behavior over time. This leads to the introduction of concepts such as overbooking and protecting seats in each ticket class. We attempted to include as many of these ideas into our model. While our model produced reasonable results, we realize that in an ever-changing industry with growing amounts of data, there is so much potential for a more developed and better-refined model to more accurately model airline ticket prices.

6 References

[1] "Dynamic pricing strategy for airlines: Exploring current problems and adopting continuous pricing." (2021), [Online]. Available: https://www.altexsoft.com/blog/dynamic-pricing-airlines/. (accessed: 11.12.2023).

[2] "How often do flight prices change in a day?" (2023), [Online]. Available: https://aerocorner.com/blog/how-often-flight-prices-change/. (accessed: 11.12.2023).

[3] K. Shepherd. "The algorithm behind plane ticket prices and how to get the best deal." (2019), [Online]. Available: https://illumin.usc.edu/the-algorithm-behind-plane-ticket-prices-and-how-to-get-the-best-deal/. (accessed: 11.12.2023).

[4] H. Postlethwait. "Flight q&a." (2017), [Online]. Available: https://intrans.iastate.edu/news/flight-q-a/#:~:text=A%202016%20survey%20from%20Airlines,of%20people%20traveled%20for%20leisure. (accessed: 11.12.2023).

[5] A. Ory and K. Williams. "Is dynamic airline pricing costing us?" (2023), [Online]. Available: https://insights.som.yale.edu/insights/is-dynamic-airline-pricing-costing-us. (accessed: 11.12.2023).

[6] J. M. Betancourt, A. Hortaçsu, A. Oery, and K. R. Williams, "Dynamic price competition: Theory and evidence from airline markets," National Bureau of Economic Research, Working Paper 30347, 2022. DOI: 10.3386/w30347. [Online]. Available: http://www.nber.org/papers/w30347.

[7] "Number of days in advance to buy plane tickets in the u.s. in 2017, by trip purpose." (2018), [Online]. Available: https://www.statista.com/statistics/862524/airline-travelers-us-how-far-travel-is-booked-in-advance-by-trip-purpose/. (accessed: 11.12.2023).

[8] OpenAirlines. "How can airlines boost flight efficiency by reducing commercial weights?" (Sep. 2021), [Online]. Available: https://blog.openairlines.com/how-can-airlines-boost-flight-efficiency-by-reducing-commercial-weights.

[9] N. Geographic, "How Cities are Transforming in the Age of COVID-19," Oct. 2021. [Online]. Available: https://www.nationalgeographic.com/environment/article/urban-expeditions-graphic-V21 (visited on 12/14/2023).

[10] H. Aerospace. "4 ways changes in aircraft design and components reduce fuel consumption." Accessed: 2023-12-14. (2021), [Online]. Available: https://aerospace.honeywell.com/us/en/about-us/blogs/4-ways-changes-in-aircraft-design-and-components-reduce-fuel-con.