Effects of price discrimination on airline ancillary good sales: a multiple treatment propensity score weighting approach.

M.F. Ambrosius*

Master thesis research Business Administration

University of Twente, Enschede, The Netherlands

November 2017

Abstract

Ticket fares have decreased by almost 40% over the last two decades, while demand has doubled. With the decreasing margins on tickets, ancillary products and services have become an important source of income for airlines. This study is motivated by the prevalence to charge for ancillary services and investigates the effects of different prices on customer purchasing behavior of checked baggage, in addition to the flight ticket. Empirical sales results from a low-cost carrier were used, in which customers were assigned to one of the four price levels, between June 2016 and June 2017. Segments were created by group size (group size 1 versus >1 segments) and number of days before departure (DBD) booked (DBD ≤3 or ≥16 versus DBD >3 or <16 segments), to get granular results that would otherwise be eliminated by larger sample noise. Multiple propensity score analysis was used to balance the six booking specific covariates. The average treatment effect of the different price treatments on the purchase decision was assed using propensity score weighting. The population was distributed over the four treatments; -2 (12%), 0 (56%), +2 (20%), and +4 (12%). The propensity scores reduced the average standard deviation by 33%, from 0.166 to 0.112. Only the two segments with group size >1 were significant at the 0.05 level. All price treatments were estimated to be insignificant at any of the segments. Probabilities to buy at the unweighted data were both higher and more differentiated than after propensity modelling, meaning the population was subject to selection bias. The research showed that customers were not significantly affected by the price treatments, and that their checked baggage purchasing probability is strongly correlated to DBD and group size.

Keywords: multiple propensity scoring; treatment effect; airline ancillaries; revenue management

1 Introduction

Airlines primary product is flight tickets. They additionally sell secondary (optional) products; ancillaries. Revenue from ancillary products, such as checked baggage, is becoming an increasingly important revenue source for airlines, yet empirical pricing research on this topic is scarce. This paper contributes by investigating the effect of different baggage fees on purchase behavior of different types of air travelers.

1.1 Ancillaries

Airline yields have deteriorated over the last decades, as price competition grew by the increasing transparency of ticket fares. Consumer willingness to pay decreases and revenue management systems observe less demand for higher fare classes (Warnock-Smith, O'Connell, et al., 2017). The traditional revenue management systems are built on selling tickets and have a hard time maximizing revenues from these ticket sales. Ticket fares have decreased by almost 40% over the last two decades. In the meantime, demand is rising even faster, as the passenger volume has doubled (Amadeus & Accenture, 2017).

In recent years airlines have begun to unbundle their offerings by introducing ancillary products (Tuzovic, Simpson, et al., 2014). These secondary services where previously included in the ticket fare, but are now separated to decrease ticket fares even more, while adding sources of income. Examples of ancillary goods are checked baggage, priority boarding, seat reservation, and on-board consumptions.

* Author correspondence. Tel.: XXXXXXXXXXX.
E-mail address: XXXXXXXXXXX

Ancillary revenue is of growing importance to the airline industry (Odegaard & Wilson, 2016). Revenue from ancillary sales almost doubled from 4.8% in 2010 to 9,1% in 2016, as a percentage of cumulative airline revenue. Top ancillary revenue performing airlines book around 40% of their revenue from ancillaries. For low-cost carriers (LCC's) the average is 11,8% (IdeaWorksCompany & CarTrawler, 2016).

With the growing ancillary revenue share, airlines have recognized the competitive value of ancillaries. This research addresses the causal effects of different ancillary prices, specified for multiple segments, on the buying behavior of travelers.

1.2 Pricing mechanisms

Having defined their ancillary good strategies, LCC's are now trying various price mechanisms to optimize revenue from ancillary goods (Navitaire, 2016). The principle is comparable with ticket fares, but still less developed. Advances in data science and retailing flexibility provide opportunities to move towards ancillary pricing models that allow to set the right prices for the right customers.

Airlines show different levels of maturity regarding ancillary pricing (Amadeus & Accenture, 2017): Many airlines still offer static price points, and thus fail to benefit from the emerging pricing opportunities. Others use basic calculations where price a function of flight leg distance or ticket fare for example, where longer flights and high fares are related to higher ancillary prices. The minority of airlines applies advanced techniques, such as integer linear programming optimization, dynamic customer grouping, and variable pricing based on price elasticity of demand per product type. Often data science outcomes are combined with heuristic knowledge airlines have about their customers, competitors, or their strategies, to set ancillary prices. LCC's moved ahead from full service carriers in the maturity pricing optimization, as LCCs' ancillary revenue is a greater share of their overall revenue.

1.3 Research contribution

The framework in this paper is based on a discrete customer choice model and uses revealed choice data. In generic terms, we observe a single seller with a fixed capacity of perishable primary goods which are sold over a limited time frame and are dynamically priced. Additionally, the sales of primary goods make secondary goods available, of which the prices are currently fixed.

Given the importance for airlines of growing revenue share from ancillary goods, this paper tries to contribute to the rapidly developing ancillary topic by investigating price treatment effects to different customer segments. To my best knowledge, this study is the first to research customer decision making for an ancillary good, specifically checked baggage, using airline price testing data. First, customer segments are defined within the LCC environment, using ticket purchase attributes. Second, multinomial propensity scoring is used to demonstrate segmented price treatment effects on purchasing behavior. Third, using historical sales data, revenue optimizing treatments are suggested.

The paper is structured as follows. In section 2, literature on multinomial propensity models, and airline ancillary revenue management are reviewed. In section 3, the data collection and variables are discussed. The section also explains the statistical method of this study. The fourth section presents the results; descriptive statistics of the sample, segmenting covariates and their cut points, propensity scoring estimates, balances, and finally the segmented price treatment effects. These results are then used for checked baggage pricing improvement conclusions. The paper ends with a conclusion, managerial implications, a discussion of limitations, and directions for future research.

2 Literature

The five-stage literature review of Wolfswinkel, Furtmueller, et al. (2013) was followed. To retrieve (most of) the papers from ScienceDirect the following keywords where used; ancillaries, baggage, fee, airline, pricing, revenue, propensity model, treatment effect. Through this iterative process papers are retrieved that to focus on either the propensity modelling method, or ancillary revenue management.

2.1 Propensity models

In dynamic pricing research, there are many scenarios where the objective is to disclose causal relationships, such as the effect of customer profile characterizing variables (further referred to as covariates) and the willingness-to-pay. Finding this information from observational studies is challenging because the observed association between "causal" variable and output variable not only results from the effect of the former on the latter, but also from confounding factors (Pearl, 2000). Confounding occurs when covariates are related to both outcome and treatment assignment, resulting in systematic differences between treated and control groups.

Potential confounding variables can be dealt with by randomizing experiments (Fisher, 1971). Randomized experiments are considered the standard for evaluating causal treatment effects, since it ensures that the observed covariate distribution is the same for all treated groups. The treatment can then directly be estimated by comparing the outcomes of the groups. However, often in non-experimental or observational studies randomizing is not always feasible, ethical or economical viable; e.g. medical treatments and court trials are highly context specific and mostly conducted once per case (Rosenbaum, 2010).

To establish causal treatment effects in observational studies, statistical techniques exist that adjust for confounding to approximate unbiased causality estimates. Propensity score analysis is often used to minimize the effects of bias and confounding. The propensity score is developed by Rosenbaum & Rubin (1983) as the conditional probability to be assigned to a treatment based on a set of subject specific covariates, and is used to 1) remove the effects of (selection) bias, 2) analyze causal effects of treatments. The technique aims to mimic randomized controlled trials, by balancing covariates of observations between the different treatment groups. Improved balance between groups is achieved by weighting, matching, or stratification of observations between treatments based on the propensity score. This technique seeks to isolate the treatment as the only difference between treatment and control groups.

This research takes an extension on the regular propensity analysis by studying the effects of multiple treatments, instead of a single treated group versus a control group. Research on multiple treatment propensity scoring is scarce, but the approach is proven to work well up to five treatment groups (Egger & von Ehrlich, 2013).

Propensity score analysis is used in a variety of areas such as social sciences, economics, and health sciences, proving its usefulness and generic applicability (Ferreira, Thomas, et al., 2017). To the best of our knowledge propensity score analysis has not been applied to airline ancillary pricing before.

2.2 Revenue management

In the last decades, fast amounts of research have been done on airline revenue management, both in industry and the academic field. Rich literature exists on capacity, demand, and pricing problems (Bitran & Caldentey, 2003; Chan, Simchi-levi, et al., 2004; Friesz, Kwon, et al., 2012; Gallego & Ryzin, 1994; Giaume & Guillou, 2004; Maglaras & Meissner, 2006; Mcafee & Te Velde, 2006; Odegaard & Wilson, 2016; Perakis & Sood, 2006; Yang, Zhang, et al., 2014). However, most of those revenue management activities focus on the primary good only; i.e. flight ticket sales. As described earlier, new opportunities arise with the unbundling of ancillary goods from tickets, whereas previously bundling was financially more attractive (McAfee, McMillan, et al., 1989). At the same time, little empirical research has been done on ancillary product pricing.

Where this paper looks at customer purchasing behavior of checked baggage, other papers look at the impact of the introduction of such ancillary products on their environment. Nicolae et al. (2016) empirically investigated whether the introduction of checked baggage fees, versus free bags, resulted in improved operational performance, such as on-time departure. On-time departure performance was found to become worse for LCCs, while it improved for FSCs. The introduction of checked baggage fees led to passengers taking more carry-on baggage, resulting in passengers struggling to fit bags into bins and the surplus has to be sent down to cargo compartments. So, the checked baggage policy altered passengers' buying behavior resulting in increased boarding time, and ultimately can lead to negative financial implications. This strengthens the relevance of this research, as the aim of increased checked baggage conversion not only increases revenue, but may increase on-time departure performance as well.

A study by Brueckner et al. (2015) investigated whether airline's ticket fares fell when it introduced checked baggage fees. They empirically found strong evidence that ticket fares fell. The results also showed that the average fare decrease was less than the introduced checked baggage fee, so that the trip price increased when a passenger chose to purchase checked baggage. Additionally, Scotti et al. (2015) assessed the impact of baggage fees on passenger demand together with ticket fares. They found that a $1 fare increase had a nine times greater impact on ticket sales compared to a $1 increase in baggage fee. Passengers thus consider ancillary fees to a lower extent as ticket fares, when they purchase tickets. In other words, the price of ancillaries is relatively inelastic compared to ticket fares. The studies by Brueckner et al. (2015) and Scotti et al. (2015) estimated simultaneous effects of primary and ancillary goods. This study looks in a more isolated way at the ancillary product and its price treatment effects. Despite the fact that the ancillary relation to ticket fare is not the purpose of this research, it is incorporated as a covariate in the propensity model.

Focusing on ancillaries solely, Warnock-Smith et al. (2017) examined the willingness to pay of passengers for different ancillary goods. They found that foods, checked baggage and seat assignment where perceived 'necessity' goods. They also found differences in willingness-to-pay for different goods based on carrier type (LCC/FSC/Charter), length of flight, and travel reason (leisure, business). The study was based on a preference based online survey rather than realized sales data, but still gives insight into segments for our research.

Research by Odegaard & Wilson (2016) indicated the relevance of this study's context. They found that the revenue gain from separate ancillary goods is highest in settings where the majority of customers does not demand ancillaries, which is the case for LCC travelers.

In summary, extensive research has been done on revenue management and ticket pricing. Little, but relevant, research has been done on airline ancillary goods in relation to tickets. The benefits of setting the right prices are proven to be both financial and operational. This paper builds on prior research, with a more empirical focus, and focusses on price treatment effects on customer checked baggage purchase behavior.

3 Materials and method

3.1 Data

Collection and treatment

This study uses anonymized data from a LCC. The data comprises 4891 sales records from between June 2016 and June 2017 on a single flight leg. The data recorded, and thus research scope, is further limited:

1. From ancillary goods, only separate checked baggage sales are included.
2. The observation data are ticket bookings, including those that do not contain checked baggage.
3. Observations are measured at the check-out page of booking a flight, and thus all include a flight ticket. Baggage purchases after direct ticket booking are excluded, as well as consumers that did not book a ticket at all.

The data result from an A/B/n price testing experiment, in which the population was assigned to either one of the three treatment groups or the control group. Individuals were offered the option to purchase checked baggage and other ancillary products after selecting a flight, on the

Table 1. Baggage prices by treatment.

Treatment (0 = control)	15 kg	20 kg	25 kg	30 kg	40 kg	50 kg
-2	€16	€20	€24	€33	€42	€70
0	€18	€23	€28	€38	€48	€78
+2	€20	€26	€32	€43	€54	€86
+4	€22	€29	€36	€48	€60	€94

Table 2. Covariates related to ticket bookings.

Covariate	Description
DBD (Days before departure)	Number of days between moment of booking and first flight departure
Group size	Number of people in booking
Ticket yield	Average ticket fare per flight movement per person in booking
Ticket bag ratio	Price of 15kg bag divided by ticket fare
Buy	Dichotomous variable describing whether checked baggage was purchased
Treatment	Name of the checked baggage price level shown to customer
DoW (Day of week first flight)	Day name of date at which (first) flight departs
LoS (Length of stay)	Number of days between first and second flight, if two way
Two way	Dichotomous variable describing whether two flights are booked

website of the airline. In the non-treatment situation individuals are exposed to the base price of the checked baggage. The treatments consist of three price groups, in which the prices of the six baggage alternatives presented change in proportion. The name of the treatment group refers to the price change of the 15kg bag for each specific treatment relative to the non-treatment price. A full product pricing overview is given in Table 1.

There were no missing values in the observed data, since website log-data was used and all fields for flight and baggage selection are required fields. As an exception, the length of stay was calculated, and only has a value if a two-way ticket was booked.

Independent variables

In addition to the treatment variable, latent variables are created from the raw dataset that are interpretable to both model and people. The latent variables are based on measure availability in the pricing experiment, ticket pricing literature (Bitran & Caldentey, 2003; Fedorco & Hospodka, 2013) – since studies on dynamically pricing of airline ancillary goods is scarce – and by business insights (Amadeus & Accenture, 2017). The latent variables are contextual variables that characterize each booking profile and on which the propensity score is based. ‘Ticket bag ratio’, ‘Length of stay’, and ‘Two way’ are the latent variables. Customer demographic information was not yet specified at the point where the checked baggage alternatives are displayed, and is thus not used as covariates. Although it may be interesting for research purposes. The observed and calculated covariates are explained in Table 2.

Dependent variable

The dependent or output variable “Buy” is a dichotomous factor variable. “1” means there is at least one piece of checked baggage in the booking. “0” means no checked baggage was purchased. No distinction was made between the number of bags and which bag was purchased. Problem complexity was reduced to focus on the multiple treatment effects on the granular level of bag purchase versus no bag purchase, rather than study which bag one purchases. From a business perspective, the main financial interest is in increasing the revenue from checked baggage sales.

Segmentation

Multiple segmentation is used in the construction of customer segments. Segmenting may lead to deeper understanding of - and more homogenous subsets, if there is significant difference in the dependent variable between segments. There may be different correlations for the same covariates at different customer segments, which may be covered up by larger sample noise or weaker correlations within larger groups. I.e. there may be an improved prediction on 2 segments which is being diluted by a 3rd segment when they are not recognized. The covariates are used to identify change in checked baggage purchasing behavior. To identify the covariates to segment on, a multinomial logistic regression on the dataset was run, with “Buy” as dependent variable, excluding the treatment. The decision of which covariate to use for segmentation was made by the height of the estimates and significance. Not more than two covariates were chosen. Each additional covariate dimension increases the number of segments exponentially, and

Table 3. Descriptive statistics by variable class.

Numeric Variables

Variable	Type	Complete	n	Mean	sd	Min	25% quantile	Median	75% quantile	Max	Histogram
DBD	integer	4891	4891	46,63	35,39	0	17	39	70	184	
Groupsize	integer	4891	4891	1,64	1,11	1	1	1	2	24	
TicketBagRatio	numeric	4891	4891	0,43	0,29	0,07	0,22	0,34	0,54	1,69	
TicketYield	numeric	4891	4891	63,87	35,07	10,97	38,02	58,02	83,02	265,08	
LoS	integer	2409	4891	4,61	4,22	1	3	4	5	88	

Character Variables

Variable	Type	Complete	n	n_unique	Min length	Max length
BookingID	character	4891	4891	4891	8	8
ProductID	character	697	4891	6	14	14

Factor Variables

Variable	Type	Complete	n	n_unique	Shares						
Buy	factor	4891	4891	2	0 (86%)	1 (14%)					
Treatment	factor	4891	4891	4	-2 (12%)	0 (56%)	+2 (20%)	+4 (12%)			
DoWfirstflight	factor	4891	4891	7	Sun (12%)	Mon (10%)	Tues (17%)	Wed (14%)	Thurs (19%)	Fri (19%)	Sat (8%)
Twoway	factor	4891	4891	2	0 (49%)	1 (51%)					

decreases sample size per segment. It may result in lower statistical quality and an increasingly complex pricing framework to apply in practice. Logistic regressions on single covariates was run to estimate segmented relationships by the slope parameters and contrast points where the relation (slope) changes. The maximum number of contrast points was set to three. To our knowledge, research on contrast points is not available in this context, and limited in general. Segments only add value if there is a significant difference in output between the segments. Output means and p-values were calculated to estimate the differences.

3.2 Statistical analysis

In the previous section the set of covariates was specified that potentially affect both the causal variable (treatment) - and outcome variable (purchase decision). The analysis consists of two steps. First, descriptive statistics where calculated to get rough understanding of the data. Second, the propensity score analysis was used to estimate how the observed results were affected by bias, and to disclose causal effects of the treatments on the purchasing output. A recommended and frequently used five steps method was followed, which is based on existing research and combines a counterfactual framework for causal inference, explicit causal contrast study design, and propensity scoring methods into accessible procedures (Li, Wen, et al., 2017).

This research observes three treatments versus one control group. Therefore, an extension on the regular propensity scoring model was used. It means that "normal" propensity scoring was done for each possible treatment duo combination; '-2' to '0', '+2' to '0', '+4' to '0', '-2' to '+2', etc.. Next the effects of each treatment duo are consolidated into single estimates per treatment.

To do so, first, a propensity score for each observation (in this analysis, each booking) was estimated from a regressions model predicting the likelihood being assignment to each of the treatment groups, based on the booking specific covariates. The propensity score thus results from the values of the covariates.

Second, after propensity score estimation, the propensity score distribution was compared between the different treatments (including control group), to check whether there are systematic treatment assignment differences within the population. Systematic differences are an indicator for confounding, for which should be corrected.

In step three, different corrective matching techniques can be applied, such as matching or weighting. These techniques correct for selection bias and confounding effects by only comparing bookings with similar covariate values (propensity scores), but given different treatments. Propensity score 'weighting' was preferred over the more frequently used 'matching' method for the following reasons: 1) matching throws away unmatched observations. The matched sample may therefore not be representative for all customer groups 2) matching may lead to few matches in high dimensional covariate spaces, 3) matching may lead to few matches when searching for similar cases through all four treatment populations, 4) matching is basically a limited version of weighting with binary weights.

Fourth, a balance check was done to make sure that systematic differences between the treatments were reduced. Standardized differences were estimated for each covariate, and should be more equal after the weighting process than before.

The fifth and final step simulated the average treatment effects (ATE), at the population level, of giving the entire population within a segment a specific treatment. The alternative is to estimate the average treatment effect on the treated populations only (ATT). The interest is in finding optimal price treatments for the entire segment population, with respect to no treatment. The ATE was estimated as the mean difference between treatment and non-treatment population outcomes. Consequently t-statistics were calculated and treatment effects were estimated including p-values. Since the problem is a logistic problem, the output estimates are log transformed

Table 4. Logistic regression results of input contribution to consumer purchase behavior.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-3,14	0,27	-11,69	< 2e-16
Groupsize	7,96	1,22	6,54	0,00
TicketYield	0,97	0,66	1,46	0,14
TicketBagRatio	-0,65	0,47	-1,38	0,17
DBD	1,68	0,34	4,99	0,00
LoS	0,12	0,01	8,00	0,00
TicketBagRatio	-0,65	0,47	-1,38	0,17
DoWfirstflight S	-0,16	0,18	-0,92	0,36
DoWfirstflight M	0,50	0,17	2,97	0,00
DoWfirstflight T	0,49	0,17	2,92	0,00
DoWfirstflight W	0,41	0,18	2,26	0,02
DoWfirstflight T	-0,13	0,16	-0,82	0,41
DoWfirstflight F	0,26	0,16	1,64	0,10

(log odds). The log odds are log ratios of favorable outcomes to unfavorable outcomes, i.e. the chance that an event will occur divided by the chance that it won't (Fulton, Mendez, et al., 2012). It can be transformed to a probability interpretation, in which probability is defined as the fraction of desired outcomes in the context of every possible outcome (Fulton, Mendez, et al., 2012). In a mathematical form:

Logistic function,

$$\log(odds) = \beta_0 + \beta_1 X_1 + \ldots + \beta_n X_n$$

And,

$$odds = e^{\beta_0 + \beta_1 X_1 + \ldots + \beta_n X_n} = \frac{probability}{1 - probability}$$

Then,

$$probability = \frac{odds}{1 + odds} = \frac{e^{\beta_0 + \beta_1 X_1 + \ldots + \beta_n X_n}}{1 + e^{\beta_0 + \beta_1 X_1 + \ldots + \beta_n X_n}}$$

The regression and multinomial propensity score estimation were conducted in R, using the Twang package.

4 Results

4.1 Descriptive statistics

Descriptive information on the booking data was presented in Table 3. The checked baggage conversion rate ("Buy") is 14%. 86% of the bookings does not contain additional purchased baggage. The average traveler pays €64 for a ticket, books around 47 days before departure, and stays a slightly less than five days at the destination in case a return ticket was purchased. The average group size is 1,64. The number of product ID's is 6, which equals the number of baggage weight alternatives offered. Remarkable is that only half of the customers purchase a two-way ticket. The treatments were not equally divided over the observations. 56% of the treated was part of the control group, 20% got the '+2' treatment, and the '-2' or '+4' group both had 12% assigned.

Table 5. Customer segments

		Days before departure	
		x ≤ 3 or x ≥ 16	3 < x < 16
Groupsize	1	1* | 178** | 0,090***	2* | 656** | 0,056***
	>1	3* | 3816** | 0,156***	4* | 241** | 0,207***

*Segment **Observations ***Unweighted mean of Buy

4.2 Segments

In the first step of the analysis, the likelihood to buy was regressed on the set of covariates, excluding the treatment variable (Table 4). The logistic regression on the data shows that Groupsize, DBD, and LoS significantly contribute to the consumer purchase behavior, of which the first two have the highest impact on the output. The input data was standardized. Log transformation did not result in improved results for any of the variables. Each coefficient estimate is the change in log odds comparing the standardized inputs. Contrast point estimation on Groupsize and DBD gave covariate values to create segments. The segment numbers and sizes are shown in Table 5. The differences in means of the outcome between the segments suggests that the populations have different purchase behavior. The likelihood to buy checked baggage is less for customers that book between three and sixteen days before departure than for customers booking early or in the last three days. Also, the group size seems to affect customer likelihood to buy. It should be noted that these numbers are not yet corrected for confounding and selection bias. Using segments is helpful to disclose different correlations in the different groups, that would otherwise be covered up by larger sample noise or weaker correlations.

4.3 Propensity score analysis

Covariate differences between treatments should be minimized to estimate unbiased treatment effects.

Table 6. Covariate means per treatment

Treatment	DBD	Grpsize	TicketYld	TicketBagRatio	LoS
-2	47,34	1,70	64,21	0,36	4,38
0	49,62	1,72	61,95	0,43	4,82
2	36,09	1,38	69,07	0,43	4,39
4	49,22	1,63	64,15	0,51	4,13

Covariate means were calculated to roughly check for differences between treatments without segment discrimination (Table 6). The covariate differences show that treatments may not have been randomly assigned, or at least do not behave so due to structural confounding. Therefore, eliminating bias from the covariates using propensity modelling before estimating treatment effects will add value to this study. Not too much attention should be paid to assess covariate significance when calculating propensity scores for each observation to have received each of the treatments, because at this point they are only used to calculate propensity scores (Li, Wen, et al., 2017). Therefore, all seven covariates were included to obtain the propensity scores.

The multinomial propensity model relies on tree-based regression models that are built in an iterative fashion. The model becomes increasingly complex as the number of iterations grows or regression trees are added to the model. It is important that the models are run for a sufficient large number of iterations, to find the lowest absolute standardized mean difference (ASMD, also referred to as the effect size). The optimum number of iterations was found around 200 while 3000 iterations were run (Figure 1). So, we have evidence that the model does not have to be re-run with more iterations.

A main assumption for propensity scoring is that each observation has a non-zero probability for receiving each of the treatments. This assumption was tested by examination of the overlap of the propensity score distributions between treatments, at the previously defined iteration optimum. The overlap between

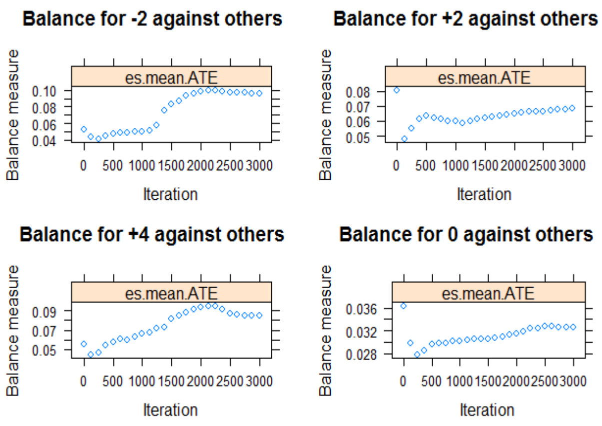

Figure 1. Balance statistic of treatments against others (absolute standardized mean difference).

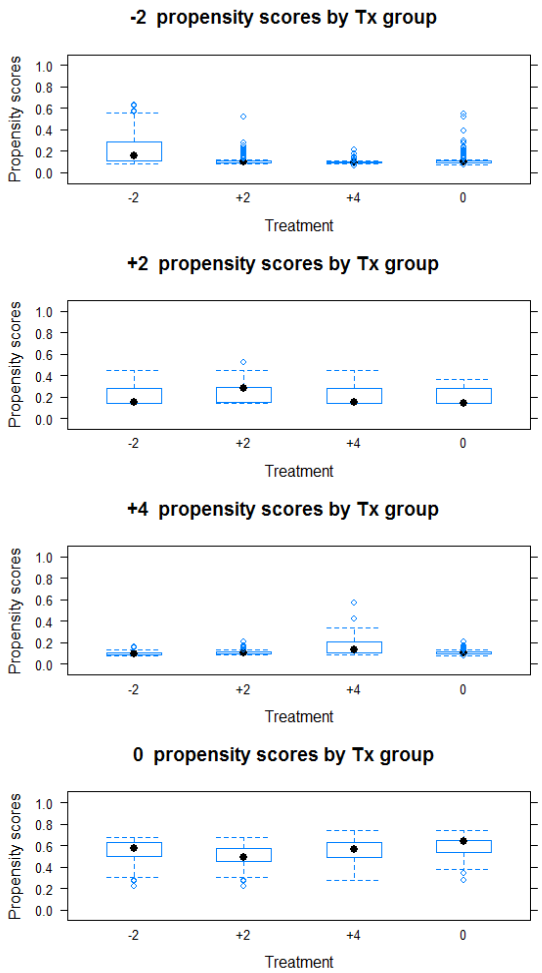

Figure 2. Overlap of the propensity score distributions.

treatment groups for each segment is shown in Figure 2. The sizeable overlap indicates that the observations in the treatment groups are comparable; i.e. the overlap assumption was met. Although, there can be some concern that in the third plot the observations with '+4' conditions do not match well with the '-2', '0', and '+2' groups. The same applies to the '-2' treatment in the top plot. There probably has been some selection bias in giving the '-2' and '+4' treatment.

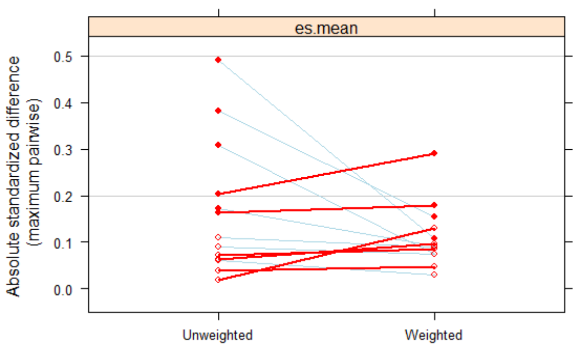

Figure 3. Comparisons of the absolute standardized mean differences on the pretreatment covariates, before and after weighting.

Next, including normalized propensity weights to each observation led to significant covariate difference (bias) reduction. The balancing test in Figure 3 shows how adding the weights led to lower overall absolute standardized differences. The thick red lines show that for some covariates the pairwise absolute standardized differences increased after weighting. Overall the differences decreased, which means that there were enough booking profiles with different treatments to compare. Better balanced set of covariates equals a less biased comparison between the populations and how the treatments affected their purchasing behavior.

The summary statements of the segmented propensity models showed that the effective sample sizes of the weighted set are similar to the original sample size (Table 7). Thus, little data was omitted because of the weighting and the weighted samples set therefore is representative for the originally observed population.

4.4 Treatment effects

After segmenting, propensity scoring, weighting, and balance assessment, the estimates of the average treatment effects (ATE) where conducted on the weighted data set. Table 8 shows the logistic ATE

Table 7. Sample sizes and effective sample sizes.

Segment	Treatment	Orig. sample sz	Effec. sample sz
1	-2	12	11,7
	0	80	80,0
	2	60	60,0
	4	26	22,7
2	-2	70	69,4
	0	321	292,7
	2	202	169,3
	4	63	42,0
3	-2	467	285,9
	0	2239	2200,9
	2	679	553,2
	4	431	246,8
4	-2	26	18,9
	0	160	157,0
	2	25	24,1
	4	30	27,4

Table 8. Estimates from propensity weighted multilevel logistic regression predicting checked baggage purchase decision.

		Treatments				Treatments and covariates			
Segment	**Covariate**	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**
1	(Intercept)	-1,63	0,77	-2,09	0,04	1,70	1,95	0,87	0,38
	Treatment 0	-0,89	0,89	-1,00	0,31	-1,56	1,30	-1,21	0,23
	Treatment +2	-0,77	0,91	-0.85	0,39	-1,45	1,28	-1,13	0,26
	Treatment +4	-0,46	1,01	-0,46	0,64	-1,13	1,30	-0,87	0,39
	TicketYield					-2,28	3,52	-0,65	0,52
	TicketBagRatio					-6,44	9,83	-0,66	0,51
	DBD					-204,50	63,06	-3,24	0,00
	DoWfirstflight S					-1,87	0,81	-2,30	0,02
	DoWfirstflight M					-0,01	0,99	-0,01	0,99
	DoWfirstflight T					-0,38	0,82	-0,47	0,64
	DoWfirstflight W					0,07	0,89	0,08	0,94
	DoWfirstflight T					1,39	0,78	1,77	0,08
	DoWfirstflight F					-0,72	0,94	-0,77	0,45
	Covariate	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**
2	(Intercept)	-3,16	0,59	-5,32	0,00	-2,19	1,17	-1,87	0,06
	Treatment 0	0,04	0,66	0,06	0,95	-0,02	0,66	-0,03	0,98
	Treatment +2	0,60	0,65	0,92	0,36	0,55	0,66	0,84	0,40
	Treatment +4	0,00	0,85	0,00	1,00	-0,02	0,81	-0,02	0,98
	TicketYield					0,02	2,05	0,01	0,99
	TicketBagRatio					-0,66	2,48	-0,27	0,79
	DBD					-22,37	12,04	-1,86	0,06
	DoWfirstflight S					0,25	0,56	0,45	0,66
	DoWfirstflight M					-0,88	0,62	-1,41	0,16
	DoWfirstflight T					1,13	0,59	1,91	0,06
	DoWfirstflight W					0,19	0,54	0,35	0,73
	DoWfirstflight T					-0,27	0,58	-0,46	0,65
	DoWfirstflight F					0,24	0,61	0,39	0,70
	Covariate	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**
3	(Intercept)	-1,68	0,17	-10,10	<2e-16	-2,31	0,30	-7,69	0,00
	Treatment 0	0,10	0,18	0,57	0,57	0,03	0,18	0,17	0,86
	Treatment +2	-0,24	0,21	-1,12	0,26	-0,31	0,23	-1,37	0,17
	Treatment +4	-0,22	0,24	-0,91	0,37	-0,33	0,25	-1,29	0,20
	Groupsize					7,25	2,37	3,06	0,00
	TicketYield					0,28	0,79	0,35	0,72
	TicketBagRatio					-0,22	0,59	-0,36	0,72
	DBD					1,35	0,34	3,96	0,00
	DoWfirstflight S					-0,08	0,18	-0,47	0,64
	DoWfirstflight M					0,33	0,17	1,96	0,05
	DoWfirstflight T					0,24	0,17	1,42	0,16
	DoWfirstflight W					0,16	0,18	0,90	0,37
	DoWfirstflight T					-0,14	0,16	-0,86	0,39
	DoWfirstflight F					0,10	0,16	0,59	0,55
	Covariate	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**
4	(Intercept)	-2,10	0,57	-3,67	0,00	-2,57	1,14	-2,25	0,03
	Treatment 0	0,90	0,60	1,49	0,14	0,84	0,58	1,44	0,15
	Treatment +2	0,56	0,80	0,70	0,48	0,21	0,83	0,25	0,80
	Treatment +4	0,22	0,79	0,28	0,78	-0,02	0,77	-0,03	0,98
	Groupsize					13,69	4,03	3,39	0,00
	TicketYield					0,90	1,87	0,48	0,63
	TicketBagRatio					2,05	3,16	0,65	0,52
	DBD					-16,57	10,99	-1,51	0,13
	DoWfirstflight S					0,24	0,45	0,52	0,60
	DoWfirstflight M					1,20	0,44	2,72	0,01
	DoWfirstflight T					0,60	0,45	1,33	0,19
	DoWfirstflight W					0,71	0,46	1,53	0,13
	DoWfirstflight T					0,39	0,47	0,84	0,40
	DoWfirstflight F					0,25	0,48	0,52	0,60

estimates for the control group and the three treatments, in which the purchase decision is the outcome. The '-2' treatment was automatically taken as holdout group, because the label comes first alphabetically. This is the intercept. First, only the treatment variable was used for estimation. The intercept indicates that the segment populations show contrasting purchase behavior, which are significant at the 0.001 level, except for segment 1. The treatment effects are small and close to each other within all segments. For most treatments, the estimated standard error is larger than the estimate differences between treatments, meaning that the effects overlap. Additionally, none of the treatments was found to have a statistically significant effect. The effect size estimates seem to be inconsistent; segment 2 and 4 show increasing purchase estimates for increasing prices, where treatments effects in the other segments make more sense and do the opposite.

Secondly, to improve the p-values of the treatment estimates, the ATE was calculated again with inclusion of the weighted booking covariates (the 'Treatments and covariates' column of Table 8). First, the intercepts of segment 2, 3, and 4 still showed statistically significant differences. The intercept of segment 1 has no statistical value, with a p-value of 0.8. Secondly, the treatment results are comparable to the previous regression; apart from segment 1 the estimates decreased equally. Still, none of the ATE treatment estimates was statistically significant at any of the segments. Thirdly, it also shows that the estimates of most of the added covariates are at least equal in size to the treatment estimates. The DBD and Groupsize estimates are structurally larger than the treatment estimates, and are statistically significant. In segment 1 the DBD estimate has a remarkable large value of -200 and is significant at the 0.01 level, while the intercept had no explanatory value. The findings for these segments suggested that customers do have different checked baggage purchasing behavior, but where not significantly affected by any of the price treatments, and most of the behavioral differences were caused by DBD and Groupsize.

Table 9. Effects and revenue per treatment and segment.

Segment	**Treatment**	**Odds**	**Oddsratio**	**Probability**	**MeanAncRev**	**MeanAncYield**
1	-2	0,703	1,000	0,413	18,000	7,433
	0	0,093	0,132	0,085	23,833	2,028
	+2	0,106	0,150	0,095	34,000	3,245
	+4	0,154	0,218	0,133	31,333	4,173
2	-2	0,112	0,132	0,101	20,000	2,022
	0	0,111	0,983	0,100	23,929	2,381
	+2	0,195	1,738	0,163	24,375	3,985
	+4	0,110	0,982	0,099	30,750	3,057
3	-2	0,099	1,000	0,090	25,743	2,323
	0	0,102	1,032	0,093	26,338	2,445
	+2	0,073	0,734	0,068	29,324	1,991
	+4	0,072	0,721	0,067	34,127	2,278
4	-2	0,077	1,000	0,071	27,250	1,943
	0	0,178	2,320	0,151	33,158	5,014
	+2	0,095	1,235	0,087	33,000	2,858
	+4	0,075	0,979	0,070	27,250	1,905

Next, the log odds were transposed to probabilities for easier interpretation of customer purchasing behavior differences, for which the calculations from section 3.2 were used. Although the treatment effects were insignificant, the logistic regression model estimates for the odds, odds ratios, and probability of checked baggage purchase were calculated (Table 9). The MeanAncRev gives the average checked baggage spend per booking from the bookings that did include checked baggage. The MeanAncYield gives the average checked baggage revenue per booking, measured over the total population, regardless of whether checked baggage was bought or not. The MeanAncRev follows the price changes of the treatments in most cases. Exceptions are in segment 1 and 4 where the +4 treatments generate less revenue per checked baggage purchase. This could be explained by customers picking lower weight baggage options on average, because the price of the previously preferred option exceeds their willingness-to-pay level. MeanAncRev does not indicate pricing optima for the overall revenue point of view, because conversion rates were not taken into account. The Probability variable represents the so-called conversion rates. The only significant intercepts, from segment 2, 3, and 4, suggested that the probability to buy ranged between 0.071 and 0.101. The observed means of buy in the raw data in Table 5 show values between 0.056 and 0.207 for the same segments. Thus, selection bias may have played a role in the higher contrasting price treatment effects as observed in the raw data. Lastly, MeanAncYield was calculated by multiplying MeanAncRev and probability. The higher the yield the higher the overall revenue from checked baggage sales.

5 Conclusion and discussion

5.1 Conclusion

Ancillary products and services have become an important source of income for airlines. The revenue management activities around this topic are growing, as price pressure on tickets reduces profit margins from ticket sales. To drive ancillary revenue, unbundling of services such as seat reservation and checked baggage has become common practice (Tuzovic, Simpson, et al., 2014). Revenue management for tickets is mature, whereas it is not for ancillary sales. Motivated by the lack of price effect knowledge on ancillaries, this paper analyzed empirical test results of giving customers different price treatments for checked baggage. In general terms, the case concerns a single seller, with a fixed inventory of primary goods, which additionally offers secondary goods. The main research question of the paper, and knowledge gap of the seller, was how customer behavior of purchasing checked baggage was affected by different price treatments. The typical goal of an airline would be to set ancillary prices that lead to maximized revenue.

The goal of this study was to evaluate the relationship between the given price treatments and purchase behavior. To do so, sources of treatment selection bias were accounted for. The modelling was done for four customer segments, with the assumption that contrasting correlations can be covered up by larger sample noise. Using ticket booking data only, the number of days before departure booked and the group size were identified as most contributing factors to explain the customer purchase decision, using multiple logistic regression. Contrast point were used to split the variables in parts that had different relationships to the output. For group size 1 versus >1 segments were identified. For days before departure ≤3 or ≥16 versus >3 or <16 segments were identified. Next, propensity modelling was done to reduce selection bias in treatment assignment. Propensity scores weighting kept 87% (4262 records) of the data, thus the bias corrected data set was quite representative for the observed data. The weighted propensity score values should be virtually the same for observations between the treatments. Balancing tests indeed confirmed that covariate standard deviations between treatments were successfully reduced by the weighting. The average standard deviation dropped 33%, from 0.166 to 0.112.

The treatment effects were estimated per segment by the simulation of giving an entire segment population the same treatment. This way the revenue maximizing treatment can be found per segment. The estimated differences in probability to purchase checked baggage where only significant for segment 3 and 4. Segment 1 and 2 had p-values of 0.38 and 0.06 respectively. Probabilities from the observed population were both higher and more differentiated than the causal effect inferred with the propensity modelling approach. If there was no confounding, the observed and propensity modelled data would yield the same results. The result showed that the observed population was subject to selection bias; i.e. non-randomized treatment assignment. Inferring the treatment effects from direct comparison between price treatments would have led to biased results, highlighting the importance of correcting for confounding (Rosenbaum & Rubin, 1983). Another finding of this study is that there are significant behavioral differences for travelers traveling with a group size above 1, depending on the number of days before departure the booking is made. Bookings between 3 and 16 days before departure are 29% more likely to convert to a checked baggage purchase than bookings made between 0 to 3 or more than 16 days before departure. On the other hand, it remains inconclusive whether any of the tested price treatments influenced customer checked baggage purchasing behavior. The price treatment effects where estimated to be insignificant at each of the segments. We cannot claim the treatments had no effect, because in that case the treatment estimates should have been close to zero and statistically significant. Price treatment optima can only be found if statistically significant treatment estimates are present, and then are calculated to probabilities and mean ancillary yield per booking. It is also important to stress that effects by group size and booking days before departure where significant confounders with larger estimates than any other covariate, even after partially accounting for its effects by segmenting. Additionally, the ratio between ticket price and checked baggage price (TicketBagRatio) was expected to be of more influence on customer purchasing behavior because customers are strategic buyers and sensitive to (relative) prices (Yang, Zhang, et al., 2014).

Nonetheless, it could be argued that a revenue optimization opportunity has been identified. Booking profiles with different purchasing behavior have been identified, along with differences in money spend on checked baggage. Price optimization remains a topic for further experimentation and research.

5.2 Discussion

The study presented has certain strengths. First, the studied sample was retrieved from real-world practice. No simulations or assumptions were done to retrieve the data. The data represented purchase decisions made by real customers, rather than consumer preferences from surveys. However, it led to empirical results that are not truly randomized and might be affected by unobserved confounding, which may distort findings. Secondly, the study used a multiple-propensity weighting approach to compare more than a – more often used – single treatment versus a control group. Estimating multiple effects in a single test is more efficient than running sequential tests for each treatment, and prevents environmental factors to change between tests. Thirdly, the main difference between (logistic) regression and propensity modelling is that the regression controls for covariate differences in a linear fashion. Propensity score matching or weighting eliminates the linearity assumption by repetitively estimating effects between similar cases only.

On the contrary, the interpretation of the results should be made with knowledge of a few limitations. First, if any confounding factor is unobserved, then imbalances may exist between the treatment groups at the segments. The weakness of propensity modelling mainly stands from not controlling for unobserved confounders. No upfront information was available about possible selection bias, although the treatments had non-equal number of observations. The covariate means differed as well. Therefore, there is no guarantee that the results were unaffected by unmeasured confounding. Secondly, propensity scoring is perceived as a powerful method to balance observed covariates to obtain rational estimations between treatment groups when randomization was not done or was not possible. However, it is not a substitute for randomization, and should not be interpreted as such. Propensity modelling should serve as complementary method for tests with populations that already have comparable covariate values. Which, to some extent, applies to the data in this study. Lastly, covariate selection was strictly limited by information available at the point where checked baggage prices are presented to the customer. It is thus likely that confounding effects are present in unobserved (e.g. demographic) factors.

5.3 Future research

This study has room for research extensions in a few ways. First, causal inference is extremely difficult, and it is close to impossible to control for all (hidden) confounders that bias the treatments. Still, in not perfectly randomized tests, a causal inference method like propensity modelling could be used as hypothesis generator to be validated later using randomized tests. In many applications, randomization is expensive or not possible (e.g. medical treatments, court trials), but it is for online sales of an ancillary product. Truly randomized treatment data also saves inaccuracies caused by the transformational processing steps in the case of propensity modelling. When confounding effects are the same between treatment groups, fitting simple (logistic) regression could lead to similar and even better results as propensity modelled estimations. Secondly, in addition to random treatment assignment, inclusion of more confounding factors, such as customer demographic information, might reduce bias even more and eventually improve estimation of price treatment effects. Thirdly, the lack of significance of the treatment effects within the observed context could be solved by either extending the pricing test to collect more observations, or by assigning larger price differences to treatments to push boundaries of consumer willingness-to-pay, and thereby realizing more differentiated treatment estimates. Fourth, this research solely studied price effects on checked baggage. For broader impact analysis, future research should be done on the indirect effects of price optimization for the checked baggage on other ancillary products. Revenue increase of one product could be cannibalized revenue from another product. Also, ancillary pricing was found to have effects on the ticket conversion (Scotti & Dresner, 2015). As last point for further research, what this study did not address was ancillary pricing optimization from the perspective of cumulative revenue maximization from both primary and secondary products.

5.4 Managerial implications

It depends on a firm's strategy which type of optimization is desired. The segmenting and price treatment findings can be valuable in various ways. When improving customer satisfaction is the goal, one should opt for increased conversion (probability to buy) of the ancillary products (Scotti, Dresner, et al., 2016). For direct revenue improvement, the KPI to focus on is the baggage yield, which is calculated by multiplying the probability to buy with the average spend on checked baggage. Direct revenue optimization is the most common form of optimization. On the other hand, an indirect effect could be the reduction of operational costs. Lower ancillary pricing increases conversion, which reduces the amount of carry-on baggage (Nicolae, Arikan, et al., 2016). Decrease of carry-on baggage was found to decrease passenger boarding time, and with that prevents the expenses that come with delayed flights (Nicolae, Arikan, et al., 2016). On the contrary, increased number of checked baggage also goes hand in hand with increased airport handling costs and increased risk of lost baggage (Scotti, Dresner, et al., 2016). This study aimed at estimating treatment effects only. Using the knowledge from these estimations some type of these optimizations could be fulfilled. The meaning given to the price treatment optimization depends on the strategic goals of the party applying the price discrimination. This study contributes by highlighting pricing opportunities.

The suggestions are left for future research and eventually business implementation. Hopefully, the proposed results motivate airlines to do more ancillary price experiments. Whereas the suggested modelling method hopefully stimulates for more research to be done on revenue management and ancillary products, as it is the future of airline revenue management.

6 References

Amadeus, & Accenture. (2017). *Merchandising '17: Trends in Airline Ancillaries*.

Bitran, G., & Caldentey, R. (2003). An Overview of Pricing Models for Revenue Management. *Informs Manufacturing & Service Operations Management*, *5*(3), 203–229.

Brueckner, J. K., Lee, D. N., Picard, P. M., & Singer, E. (2015). Product Unbundling in the Travel Industry: The Economics of Airline Bag Fees. *Journal of Economics and Management Strategy*, *24*(3), 457–484. https://doi.org/10.1111/jems.12106

Chan, L. M. A., Simchi-levi, D., & Swann, J. L. (2004). Coordination of pricing and inventory decisions: a survey and classification. In *Handbook of quantitative supply chain analysis* (pp. 335–392). Springer US.

Egger, P. H., & von Ehrlich, M. (2013). Generalized propensity scores for multiple continuous treatment variables. *Economics Letters*, *119*(1), 32–34. https://doi.org/10.1016/j.econlet.2013.01.006

Fedorco, Ľ., & Hospodka, J. (2013). Airline pricing strategies in European airline market, *VIII*(2).

Ferreira, V. C., Thomas, D. L., Valente, B. D., & Rosa, G. J. M. (2017). Causal effect of prolificacy on milk yield in dairy sheep using propensity score. *Journal of Dairy Science*, (August). https://doi.org/10.3168/jds.2017-12907

Fisher, S. R. A. (1971). The Design of Experiments. *Hafner Publishing Company*. https://doi.org/n/a

Friesz, T. L., Kwon, C., Il, T., Fan, L., & Yao, T. (2012). Competitive Robust Dynamic Pricing in Continuous Time with Fixed Inventories. *arXiv*, 1–39.

Fulton, L., Mendez, F., Bastian, N., & Musal, M. (2012). Confusion Between Odds and Probability , a Pandemic? *Journal of Statistics Education*, *20*(3), 1–15.

Gallego, G., & Ryzin, G. Van. (1994). Optimal Dynamic Demand Pricing over of Inventories Finite Horizons with Stochastic. *Management Science*, *40*(8).

Giaume, S., & Guillou, S. (2004). Price discrimination and concentration in European airline markets. *Journal of Air Transport Management*, *10*(5), 305–310. https://doi.org/10.1016/j.jairtraman.2004.04.002

IdeaWorksCompany, & CarTrawler. (2016). Airline ancillary revenue projected to be $59.2 billion worldwide in 2015, 1–5. Retrieved from http://www.ideaworkscompany.com/wp-content/uploads/2015/11/Press-Release-103-Global-Estimate.pdf

Li, K., Wen, M., & Henry, K. A. (2017). Ethnic density, immigrant enclaves, and Latino health risks: A propensity score matching approach. *Social Science and Medicine*, *189*, 44–52. https://doi.org/10.1016/j.socscimed.2017.07.019

Maglaras, C., & Meissner, J. (2006). Dynamic Pricing Strategies for Multi-Product Revenue Management Problems. *Manufacturing & Service Operations Management*, *8*(2), 136–148.

McAfee, P., McMillan, J., & Michael, W. (1989). Multiproduct Monopoly, Commodity Bundling, and Correlation of Values. *The Quarterly Journal of Economics*, *104*(2), 371–383.

Mcafee, R. P., & Te Velde, V. (2006). Dynamic Pricing in the Airline Industry. *Handbook Economics and Information Systems*.

Navitaire. (2016). *Three secrets to success for Ancillary Pricing Optimization*.

Nicolae, M., Arikan, M., Deshpande, V., & Ferguson, M. (2016). Do Bags Fly Free ? An Empirical Analysis of the Operational Implications of Airline Baggage Fees. *Management Science*, (March), 1–20. https://doi.org/10.1287/mnsc.2016.2500

Odegaard, F., & Wilson, J. G. (2016). Dynamic pricing of primary products and ancillary services. *European Journal of Operational Research*, *251*(2), 586–599. https://doi.org/10.1016/j.ejor.2015.11.026

Pearl, J. (2000). *Causality* (2nd ed.). https://doi.org/citeulike-article-id:3888442

Perakis, G., & Sood, A. (2006). Competitive multi-period pricing for perishable products: A robust optimization approach. *Mathematical Programming*, *107*(1–2), 295–335. https://doi.org/10.1007/s10107-005-0688-y

Rosenbaum, P. R. (2010). *Design of Observational Studies*. Springer US. https://doi.org/10.1007/978-0-387-98135-2

Rosenbaum, P. R., & Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55. https://doi.org/10.1093/biomet/70.1.41

Scotti, D., & Dresner, M. (2015). The impact of baggage fees on passenger demand on US air routes. *Transport Policy*, *43*, 4–10.

https://doi.org/10.1016/j.tranpol.2015.05.017

Scotti, D., Dresner, M., & Martini, G. (2016). Baggage fees, operational performance and customer satisfaction in the US air transport industry. *Journal of Air Transport Management*, *55*, 139–146. https://doi.org/10.1016/j.jairtraman.2016.05.006

Tuzovic, S., Simpson, M. C., Kuppelwieser, V. G., & Finsterwalder, J. (2014). From "free" to fee: Acceptability of airline ancillary fees and the effects on customer behavior. *Journal of Retailing and Consumer Services*, *21*(2), 98–107. https://doi.org/10.1016/j.jretconser.2013.09.007

Warnock-Smith, D., O'Connell, J. F., & Maleki, M. (2017). An analysis of ongoing trends in airline ancillary revenues. *Journal of Air Transport Management*, *64*, 42–54. https://doi.org/10.1016/j.jairtraman.2017.06.023

Wolfswinkel, J., Furtmueller, E., & Wilderom, C. (2013). Using grounded theory as a method for rigorously reviewing literature. *European Journal of Information Systems*, *22*(1), 45--55. https://doi.org/10.1057/ejis.2011.51

Yang, N., Zhang, R., Yang, N., & Zhang, R. (2014). Dynamic Pricing and Inventory Management Under Inventory-Dependent Demand. *Operations Research*, *62*(5), 1077–1094.