

Value of Flexibility: Evidence from Trains*

Please Do Not Distribute Without Permission!

Most Recent Version

William H. Burton[†]

November 20, 2024

Abstract

The use of dynamic pricing tools has expanded significantly over the past two decades, accompanied by a growing emphasis on offering consumers the flexibility to cancel previously purchased items. This paper leverages a comprehensive dataset comprising all sales and cancellations from a major European passenger rail company to empirically quantify the effects of cancellation options on dynamic ticket pricing, consumer welfare, and firm revenue. First, I show that the overall effect of offering this flexibility is theoretically ambiguous, as it involves both higher ticket quality and lower opportunity costs. I therefore develop and estimate a structural dynamic pricing model and conduct counterfactual simulations to explore the overall effect as well as the relative importance of the ticket quality and opportunity cost channels. My results show that removing the flexibility to cancel decreases on average prices by 1.9%, consumer welfare by 4.5%, and firm revenue by 5.7%, suggesting that both consumers and firms benefit from allowing cancellations. In my particular setting, I find a negligible impact of the opportunity cost channel, thus indicating that observed changes are fully attributable to the higher ticket quality channel. Larger changes in opportunity costs would therefore amplify the losses from restricting cancellations.

Keywords: Dynamic Pricing, Intertemporal Price Discrimination, Cancellations, Transportation Markets, Passenger Rail

JEL-Codes: L11, C61, L92, R41

*I am grateful to my advisors Jan De Loecker and Frank Verboven for their support. I would also like to thank Daniel Akerberg, Matthew Backus, Sebastian Fleitas, Mogens Fosgerau, Julian Hidalgo, Florin Maican, Marleen Marra, Anders Munk-Nielsen, Bertel Schjerning, Bernardo S. da Silveira, Jo Van Biesebroeck, and Honghao Zheng for their helpful comments and suggestions. Special thanks to MBR, SP, and TS at TC-0. Financial support from the Research Foundation Flanders (FWO) in the form of Grant #11D1524N is gratefully acknowledged. Any errors and shortcomings are my own.

[†]Department of Economics, KU Leuven, and Research Foundation Flanders (FWO)
williamharry.burton@kuleuven.be

1 Introduction

Over the past two decades, an increasing number of firms across a range of perishable good industries have adopted dynamic pricing tools to sell their existing capacity more effectively. (Boston Consulting Group, 2024). Alongside this trend, firms have also started to offer consumers the flexibility to cancel or return previously purchased items (Sampson, 2024). In the hospitality sector, for instance, it is now common for rooms to be cancellable free of charge up to the day before arrival. Transportation companies such as airlines or passenger rail services offer their passengers a similar degree of flexibility by offering rebookings and/or refunds.

The overall economic effect of providing such flexibility is, however, theoretically ambiguous. On the demand side, a consumer's willingness to pay is likely to increase, as cancellation options offer a form of insurance against unexpected changes in travel plans. On the supply side, firms may respond by raising prices to reflect the improved quality of the ticket. However, at the same time the opportunity cost of selling a seat today versus in the future is reduced, as the firm now has the ability to re-sell the same seat upon cancellation. Whether final prices rise due to improved ticket quality or fall due to lower opportunity costs, and the resulting net effect on consumer welfare and firm revenue, remains an empirical question.

This paper empirically quantifies the impact of offering consumers the flexibility to cancel on the dynamic pricing decision of the firm, and its subsequent effect on consumer welfare and firm revenue. Using ticket-level sales and cancellation data from a large European passenger rail company, I develop and estimate a structural dynamic pricing model of train-level pricing. Counterfactual simulations show that disallowing cancellations would negatively impact both the firm and consumers. In my particular setting, the results are fully driven by the higher ticket quality channel as the opportunity cost channel has a negligible impact. Larger changes in the opportunity costs would therefore amplify these losses from restricting cancellations.

Beyond the passenger rail industry's critical role in decarbonising medium- and long-distance travel, it offers several institutional advantages for studying this question. The fixed schedule of services, pre-allocated rolling stock, and physical infrastructure constraints make pricing the primary short-term mechanism to steer consumer demand. Additionally, the industry's frequent

monopolistic structure enables the clean identification of the effects of cancellations on pricing without the confounding influence of competition. Finally, revenue management systems in the passenger rail industry typically only permit upward price adjustments. This allows one to abstract from concerns about strategic consumer behaviour.

The analysis proceeds in three steps. First, I develop a theoretical dynamic pricing model to examine the trade-offs a firm encounters when offering consumers the flexibility to cancel. This model highlights the strategic considerations surrounding cancellation flexibility in pricing. Second, I present empirical evidence to inform and motivate the structural modelling assumptions. Specifically, I document substantial variation in ticket prices and sales, both across and within (origin, destination, calendar date of travel, departure time)-tuples, highlighting the need for a train-level pricing model. The data also shows that initial price levels increase with expected capacity utilisation, and that prices increase towards the departure deadline. These findings motivate capacity constraints and time-varying price sensitivities. Lastly, I approximate the revenue manager's pricing policy function by using a logistic regression of price changes on lagged sales and cancellations, alongside separate polynomial terms for days-to-departure and lagged prices. These regression results show a positive correlation between lagged sales and current price increases, as well as a negative correlation between lagged cancellations and current price increases. This highlights the need for the structural dynamic pricing model to allow for capacity recovery through cancellations.

Building on the insights from both the theoretical model and reduced-form empirical evidence, the second stage of my analysis develops and estimates a structural model of dynamic pricing at the train level. The empirical model extends the framework developed by Williams (2022) by incorporating cancellation as well as a second ticket option. In each pre-departure period, prior to any sales or cancellations, the monopolist passenger rail company sets a price that balances the immediate revenue from current sales against the value of potentially selling the same tickets in the future. After setting this price, consumers arrive and choose between two ticket types that differ in price, quality, and cancellation fees, and the outside option of not purchasing. The risk of cancellation is assumed to follow an exogenous process. This pricing decision is repeated daily until one of the following terminal conditions is met: the departure

day arrives, discounted tickets are sold out, or the revenue manager ceases discounted ticket sales early. In any of these cases, prices are automatically adjusted to predetermined levels outside the control of the revenue manager. Reflecting the reduced-form empirical findings, all model parameters are estimated at the (origin, destination, weekday of travel)-level, with arrival rates, price sensitivities, and cancellation risks allowed to vary additionally by days-to-departure. Estimation results reveal that a substantial portion of potential passengers arrives close to the departure date. This contrasts with the more typical pattern frequently observed in the airline industry, where a significant number of potential passengers tend to arrive early in the reservation period. A similarity to airlines, however, is that these late-arriving passengers display the lowest sensitivity to price.

Finally, I use the estimated demand and supply parameters to conduct counterfactual simulations. Beginning with the current situation, in which the passenger rail company permits ticket cancellations, I simulate a policy change that disallows both cancellations and the subsequent re-sale of tickets. This allows me to compute the resulting equilibrium prices, consumer welfare, and firm revenue under a no-cancellation setting. The results show that, compared to the baseline with cancellations, prohibiting cancellations and the re-selling of tickets reduces - on average - prices by 1.9%, consumer welfare by 4.5%, and firm revenue by 5.7%. To disentangle the relative contributions of the two primary channels - higher ticket quality and lower opportunity costs - I perform simulations allowing for only one of the two channels at a time. In my particular setting, I find that the opportunity cost channel has no meaningful effect, as demand intersects the opportunity cost curve at quantity levels for which changes in opportunity costs amount to less than EUR 0.01. Consequently, the observed changes are entirely attributable to the higher quality channel. This suggests that in situations with larger changes in opportunity costs, losses from not allowing for cancellations would be larger in magnitude.

Related Literature This paper contributes to three strands of literature: (1) cancellations and refunds in perishable good industries; (2) empirical studies of dynamic pricing; and (3) industrial organisation studies of the railroad industry.

Empirical studies on the role of cancellations in perishable good industries remain limited.

Scott (2022) examines the design of cancellation policies as a screening mechanism for airlines, finding that greater cancellation flexibility reduces profits more than it improves consumer welfare. The analysis presented in Scott (2022) focuses on the response of consumers to various cancellation fee structures, while assuming airline ticket prices follow an exogenous first-order Markov process. Vollmer (2024) uses college football ticket sales data to compare secondary market reselling with centralised cancellations and refunds. The results show that the flexibility to cancel raises overall welfare by 0.7% when compared to a setting of decentralised re-selling of tickets. The pricing model in Vollmer (2024) is assumed to be static, as the monopolist is required to set all prices in the first period. Similarly, Lazarev (2013) explores the impact of cancellations on ex ante optimal price paths for airlines, concluding that zero cancellation fees improve consumer welfare at the expense of firm profits. Cho et al. (2018) studies dynamic hotel room pricing in the presence of cancellations but focuses on one hotel's pricing response to competitors' price changes. This paper contributes to the literature by endogenising the dynamic pricing decisions of a capacity-constrained firm that offers the flexibility to cancel, and analysing the resulting impacts on prices, consumer welfare and firm revenue.

This paper also contributes to the empirical literature on intertemporal price discrimination and dynamic pricing. Sweeting (2012) examines whether the pricing behaviour of major league baseball ticket resellers aligns with theoretical dynamic pricing models. Hortaçsu et al. (2023) studies the organisational structure behind the dynamic pricing strategies of a large U.S. airline. D'Haultfœuille et al. (2023) evaluates the profitability of dynamic pricing relative to alternative pricing strategies using data from the French passenger rail company iDTGV. My empirical setting is similar to that of D'Haultfœuille et al. (2023), as both pricing systems require revenue managers to adhere to a pricing schedule with increasing, pre-defined price levels. In contrast, my setting allows for cancellations and I additionally benefit from observing transaction-level sales and cancellations. Lazarev (2013) explores the welfare implications of intertemporal pricing, while Williams (2022) compares the welfare effects of uniform pricing, intertemporal price discrimination, and dynamic pricing. Building on the methodology developed by Williams (2022), I extend the model to incorporate the flexibility to cancel, enabling the recovery of capacity, as well as another perishable good.

This paper also contributes to the empirical industrial organisation literature on the railroad industry. Much of this literature focuses on the US freight rail industry, with early contributions such as Porter (1984). More recent work includes Chen (2024), which analyses US freight rail data to evaluate cost and price effects driven by cost efficiencies and economies of scope resulting from rail network mergers. Degiovanni and Yang (2023) jointly examines economies of scale and scope in the same industry. In contrast, empirical studies on the passenger rail industry remain limited, with D’Haultfœuille et al. (2023) being, to my knowledge, the only other work investigating firm pricing behaviour in this industry. Given the passenger rail industry’s pivotal role in decarbonising medium- and long-distance travel, understanding the behaviour of firms in this industry is of significant policy importance.

Organisation The paper continues as follows. The institutional details surrounding the empirical setting as well as the dataset are described in Section 2. Section 3 includes summary statistics as well as the motivating evidence. The structural dynamic pricing model is presented in Section 4, and the estimation strategy is outlined in Section 5. The estimation results are presented in Section 6, followed by the counterfactual simulation results in Section 7. Section 8 concludes the paper.

2 Institutional Details & Data

2.1 Institutional Details

Train Company Zero (TC-0) offers both domestic and international passenger rail services across Europe.¹ These services encompass a broad spectrum of routes, ranging from local commuter lines to high-speed inter-city connections, each exhibiting varying levels of service frequency. The majority of TC-0’s trains make several intermediate stops along their routes, reflecting a network structure that caters to diverse travel needs. The routing and scheduling of TC-0’s trains are typically determined several months in advance, a process that requires coordination with the regulating authority to ensure efficient allocation of rail infrastructure among multiple

¹The passenger rail company, which provided the data, wishes to remain anonymous and will be referred to as Train Company Zero (TC-0) throughout this paper.

users. As is common in the rail industry, akin to the airline sector, short-term adjustments in either capacity (e.g. additional rolling stock) or service frequency (i.e. the number of scheduled services) are challenging. This rigidity arises from the pre-commitment of capacity on other routes as well as the physical constraints imposed by track and station congestion.

Potential TC-0 passengers generally have the option to choose between two types of tickets: “fixed” and “flexible”. A “fixed” ticket is restricted to a specific departure time from the origin station, thereby binding the passenger to a particular time schedule. In contrast, a “flexible” ticket permits the passenger to select any departure time on the day of travel, thus granting greater flexibility with regard to departure time. Both ticket types can be cancelled up to one day before the scheduled departure. Cancellation policies, however, differ by ticket type: holders of fixed tickets incur a cancellation fee, whereas flexible ticket holders are entitled to free cancellation. Upon cancellation, the reimbursement amount — defined as the ticket price minus any applicable fee — is refunded to the customer’s original payment method, and the ticket is subsequently marked as invalid in the system.

Tickets are generally priced dynamically based on an allotment approach with increasing, pre-defined price levels. Thus, ticket prices can therefore only remain constant or increase as the departure deadline approaches. Once the maximum price level for a specific ticket type is reached, it remains at that level until departure. The size of the allotment (i.e. the number of discounted tickets available) and the corresponding price levels are subject to variation, both within and across (origin, destination, calendar date of travel, departure hour)-tuples. This approach to ticket pricing mirrors the revenue management techniques observed in other passenger rail systems, for example, as in the context of the French high-speed inter-city passenger rail company iDTGV studied by D’Haultfoeuille et al. (2023).

European passenger rail industries are typically dominated by a single large state-owned national operator, which has historically maintained substantial market control through government subsidies, established infrastructure, and a long-standing market presence. As a result, the degree of competition within each country remains relatively limited, with only a select number of privately owned operators engaging in service provision on high-demand routes. In the context of TC-0, competition is generally confined to a single alternative operator offering similar

rail services. However, even this competition tends to be partial, as the services provided by the competing firm may differ with respect to the specific origin and destination stations served, thereby leading to only a modest degree of overlap.

2.2 Data

The primary demand-side dataset covers Train Company Zero's universe of ticket sales between 01 January 2023 and 25 September 2023. Each sales record contains a unique ticket identifier, the date and time of the purchase, the calendar date of travel, the departure time (if applicable), the particular type of ticket as well as its price. Any amenities purchased in addition to the ticket itself are also recorded at the same level of granularity. If the ticket was cancelled by the passenger, the date and time of the cancellation as well as the refunded amount are also recorded. Furthermore, a separate entry is generated when a ticket is validated onboard by a staff member.

The primary supply-side dataset contains both the actual time series of ticket prices as well as the internal pricing schedules used by TC-0 revenue managers. The pricing schedules include the price level as well as the associated number of allocated tickets for each train. By combining these data with the sales and cancellation data described previously, I am able to differentiate between manual and automatic price increases.

Additional metadata datasets are also utilised. Using a data set on the carriage composition of trains, I can recover exact measures of overall capacity. Another dataset contains the scheduled as well as actual departure and arrival times for each train. as well as actual network routing information. This allows me to differentiate trains by departure hour, and by their expected and actual travel time and number of intermediate stops.

Subset For Analysis My analysis focuses on a subset of non-stop inter-city routes serviced by TC-0². In a first step, I only select station pairs for which TC-0 is the sole operator. This allows me to abstract away from the competitive effect on prices. In a second step, I restrict the subset further to trains departing between 01 March 2023 and 25 September 2023. The

²I use the word "non-stop" to indicate that the route does not require transferring at an intermediate station. The train may make intermediate stops.

reason for using 01 March 2023 as the lower bound on time, is that this allows for a sufficient number of pre-departure sales periods to be observed. In a third and final step, I impose the restriction that flexible ticket prices remain constant across the days leading up to departure. This constraint allows me to abstract from the joint (dynamic) pricing decisions between fixed and flexible tickets, thereby simplifying the analysis considerably. The remaining subset of trains represents more than 50% of TC-0's passengers.

3 Motivating Evidence

3.1 Theoretical Insights

Before providing empirical evidence, I will first discuss the effect of allowing for cancellations through the lens of a theoretical dynamic pricing model. Without loss of generality, assume that a passenger rail company is selling exactly one seat on board its train across two periods. I evaluate the firm's pricing decision in the first period under two different cancellation scenarios. *I* will indicate a scenario in which there is no cancellation, and *II* will indicate a situation with reimbursement in the event of cancellation.

I make several simplifying assumptions in order to keep the model tractable. First, I assume that the probability of travel, μ , is exogenously given. The adverse event forcing the passenger to cancel therefore happens with probability $1 - \mu$. Second, I assume that the consumer is myopic, meaning she does not delay her purchasing decision strategically. Third, if the consumer experiences the adverse cancellation event, she will cancel her seat with probability one if she receives at least some money back. If she does not receive any money back, she will not notify the company of her decision not to travel ("no-show"). Fourth, the passenger rail company is not allowed to overbook the train. Finally, I assume that the value of still being able to sell the seat in the second period is constant and positive.

The demand side in the first period consists of a consumer making a discrete choice between buying the ticket and not buying the ticket. The expected indirect utility of buying the ticket

is given below by Equation 1.

$$\begin{aligned}\mathbb{E}[u_{i1}] + \epsilon_{i1} &= \underbrace{\mu \times (-\alpha \times p_1 + \delta)}_{\text{Travel}} + \underbrace{(1 - \mu) \times (-\alpha \times \gamma p_1)}_{\text{Cancel}} + \epsilon_{i1} \\ &= -\alpha \times p_1 \times (\gamma + \mu \times (1 - \gamma)) + \mu \times \delta + \epsilon_{i1}\end{aligned}\quad (1)$$

The first line of Equation 1 consists of two main components. In the event that she travels, the consumer must pay the full price p_1 and receives the benefits of travel, denoted by δ . If she is forced to cancel, she does not receive any benefits of travel but pays only a fraction of the original price, denoted by γp_1 . This is the cancellation fee. Rearranging terms, one can see in the second line of Equation 1, that both the probability of travel μ and the cancellation fee structure γ scale the price-sensitivity parameter α . Thus, conditional on μ , reducing the cancellation fee factor γ , i.e. offering higher reimbursement amounts, lowers the price-responsiveness of consumers.

Assuming that the unobserved vector of random utility shocks ϵ_i is independently and identically distributed (i.i.d.) across choices and follows a Type 1 Extreme Value (T1EV) distribution, and normalising the indirect utility of not buying the ticket to $u_{i0} = 0 + \epsilon_{i0}$, the demand function is given by Equation 2.

$$D_1(p_1) = \frac{\exp(\mathbb{E}[u_{i1}])}{1 + \exp(\mathbb{E}[u_{i1}])}\quad (2)$$

Since the firm can also sell its only seat in the second period, the passenger rail company faces an intertemporal trade-off when setting its first-period price. This dynamic pricing problem can be formulated as a Bellman equation. In the absence of a cancellation signal, the value function is given by Equation 3. The value function for the situation in which the cancellation signal is sent - if the cancellation event occurs - is given by Equation 4.

$$V_{1,I}(N = 1) = \max_{p_1} D_1(p_1) \times (p_1 - c_1) + \beta \times \left[(1 - D_1(p_1)) \times V_2(N = 1) \right]\quad (3)$$

$$\begin{aligned}V_{1,II}(N = 1) &= \max_{p_1} D_1(p_1) \times \left(\mu \times p_1 + (1 - \mu) \times \gamma p_1 - c_1 \right) + \beta \times \left[\right. \\ &\quad \left. (1 - D_1(p_1)) \times V_2(N = 1) + D_1(p_1) \times (1 - \mu) \times V_2(N = 1) \right]\end{aligned}\quad (4)$$

Please note that Equation 3 is nested within Equation 4 by setting $\mu = 1$. A situation of definite travel, i.e. $\mu = 1$, is therefore equivalent to a situation of no cancellation. This particular nesting follows from the assumptions that the cancellation signal is never sent if there is no reimbursement, and that the firm is not allowed to overbook the train. I therefore continue to only derive the first-order condition (FOC) for Equation 4.

$$\frac{\partial V_{1,II}(N=1)}{\partial p_1} = 0 \Leftrightarrow D_1(p_1^*) \times (\mu + (1 - \mu) \times \gamma) + \frac{\partial D_1(p_1^*)}{\partial p_1} \times \left[\mu \times p_1^* + (1 - \mu) \times \gamma p_1^* - (c_1 + \beta \times \mu \times V_2(N=1)) \right] = 0 \quad (5)$$

The optimal price in the first period, i.e. p_1^* , solves the first-order condition (FOC) given by Equation 5 by equating marginal revenue and marginal cost. By setting $\mu = 1$, one can use Equation 5 to see the FOC for Equation 3.

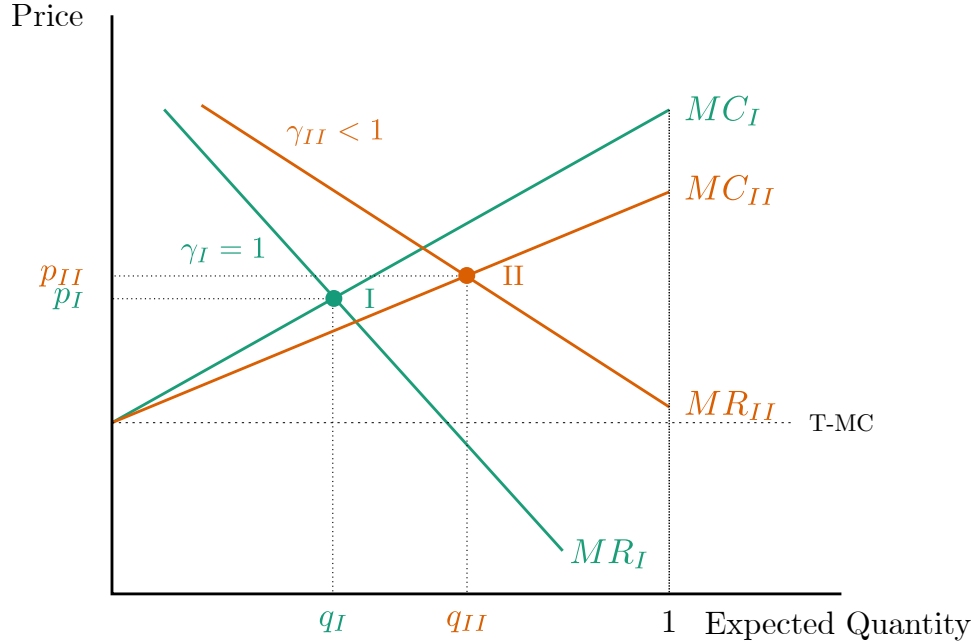
As one can see from the first order condition given by Equation 5, the marginal cost component is composed of two terms. The first term is the technical marginal cost c_1 which influences the pricing decision similarly across scenarios. The second term represents the opportunity cost of selling the seat today³, scaled by the respective probability of travel. When travel is not certain, i.e. $\mu < 1$, the opportunity cost of selling the seat in the first period is strictly lower in cases where a passenger signals their inability to travel.

Figure 1 provides a graphical illustration of this model. As indicated previously, the subscript I indicates the situation the scenario without cancellations or re-selling. In this particular setting, MR_I denotes the marginal revenue curve and MC_I denotes the marginal cost curve. At any given expected quantity level, the marginal cost comprises two components: the technical marginal cost, denoted by $T - MC$, and second, the opportunity cost. As discussed previously, the opportunity cost reflects the expected value of selling the seat tomorrow. Although the opportunity cost is not explicitly drawn in Figure 1 it is represented by the vertical distance between the respective MC line and the horizontal $T - MC$ line.

With the introduction of cancellations, two changes become apparent. First, the marginal

³The opportunity cost of selling the seat today corresponds to the value of selling it tomorrow, i.e. $V_2(N=1)$.

Figure 1: Supply and Demand in Situations Without and With Cancellation



Source: Authors' own illustrations.

Notes: Subscript I indicates the situation without cancellation, while Subscript II indicates a situation with cancellation. The marginal revenue curves are labelled MR, and the marginal cost curves are labelled MC. The horizontal line labelled T-MC denotes the technical marginal cost of providing the ticket. The variable γ indicates the share of the original price that the consumer must pay if she cancels her ticket, i.e. the cancellation fee. Please note that I have not explicitly drawn the discontinuity in MC_I , MC_{II} , $T - MC$ at $\mathbb{E}[q] = 0$.

revenue curve rotates outward, reflecting an increase in the consumer's willingness to pay. This arises from the improved quality of the ticket, as the consumer is now partially reimbursed in the event of a cancellation. This effect is depicted by the rotation from MR_I to MR_{II} , and is directly supported by Equation 1. I refer to this as the "higher ticket quality channel".

Second, the marginal cost curve rotates downwards, namely from MC_I to MC_{II} . This decrease in marginal cost arises from a reduction in the opportunity cost, as the seat is only "lost" with the probability of travel, and can be recovered and re-sold in the event of cancellation. As shown in Equation 5, the opportunity cost of selling the seat today is strictly lower for all positive values of expected quantities, and equal for an expected quantity of zero. This channel is referred to as the "lower opportunity cost channel".

The new equilibrium is characterised by a higher expected quantity, i.e. $q_{II} > q_I$, being sold a higher price, i.e. $p_{II} > p_I$, primarily driven by the outward rotation of the marginal revenue

curve. However, scenarios where the opportunity cost decreases more significantly, such as when the overall probability of travel is lower, can increase the influence of the opportunity cost channel on equilibrium prices and expected quantities. Such a case is illustrated by Figure A1 in Appendix A. While equilibrium quantities are weakly larger with any outward rotation of the marginal revenue curve and/or downward rotation of the marginal cost curve, theoretical predictions for equilibrium prices - and, by extension, consumer welfare and firm revenue - remain ambiguous. Determining the magnitude of these rotations, and their subsequent effects on prices, consumer welfare, and firm revenue, is thus an empirical question.

3.2 Summary Statistics

Table 1: Summary Statistics

	Mean	Std. Dev.	5th Pctile.	95th Pctile.
Price Ratio - Fixed / Flex	0.58	0.13	0.37	0.79
Price Ratio - Max. Fixed / Min. Flex	2.51	1.19	1.29	5.14
# Unique Fixed Ticket Prices	4.56	1.31	3.00	7.00
Share of Fixed Tickets Sold Below Max. Price	0.83	0.28	0.00	1.00
Tickets per Transaction	1.28	0.48	1.00	2.00

Source: Author's own calculations using TC-0 price, and sales data on the subset of routes described in Subsection 2.2.

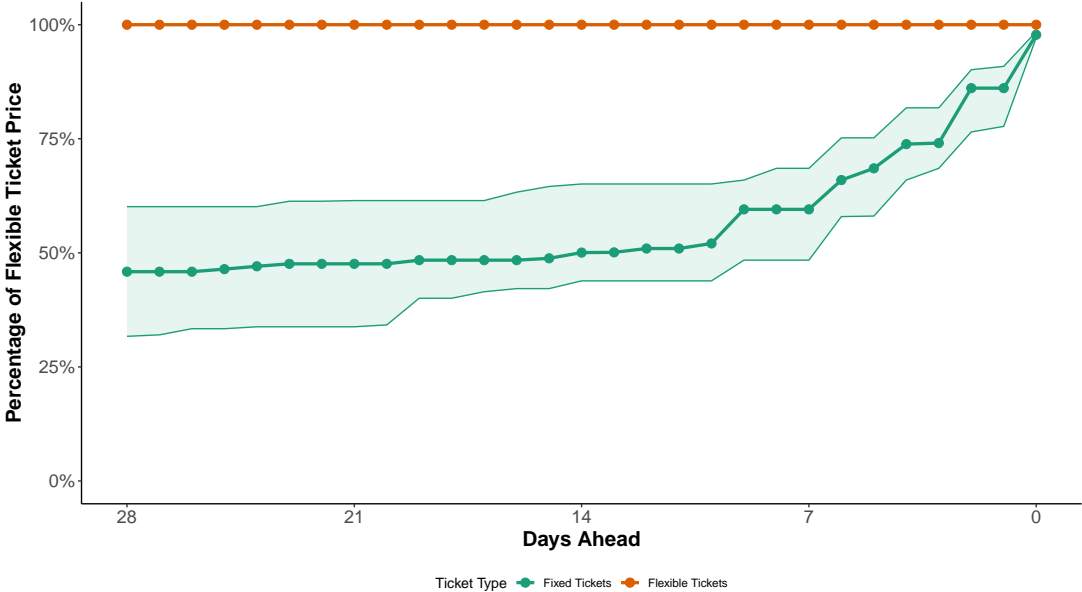
Notes: An observation is defined at the (origin, destination, calendar date of travel, departure time, day ahead)-level. The respective statistic is first computed separately for each (origin, destination, calendar date of travel, departure time)-tuple, and then the mean, standard deviation and respective percentiles are computed across those tuples.

Table 1 presents summary statistics across all trains included in the subset for analysis. On average, a fixed ticket costs about 58% of the flexible ticket. Furthermore, the variation in the price of a fixed ticket is not negligible, as the price of a fixed ticket more than doubles on average during its respective sales period. During this price increase, fixed tickets are offered at four to five different price points, corresponding to three to four price increases. Within a (origin, destination, weekday of travel)-tuple, the number of fixed and flexible tickets sold varies

substantially. The coefficients of variation⁴ for fixed ticket sales range from 0.495 to 2.261, while those for flexible ticket sales range from 0.181 to 3.659. At the same time, realised demand levels are difficult to predict. The R^2 of a regression of total fixed ticket sales on a multiplicative (origin, destination, weekday of travel, departure time)-specific fixed effect and a (calendar date of travel)-specific fixed effect is equal to 0.480. The R^2 of a regression of total flexible ticket sales on the same fixed effect specification, however omitting the departure time component from the first fixed effect, is equal to 0.804. The average number of tickets per transaction equals 1.28, with roughly 83% of fixed tickets sold below the maximum price. In its entirety, this descriptive evidence highlights significant inter- and intra-train variation in prices and sales, motivating the need to model demand and supply at the train level.

3.3 Price Changes

Figure 2: Prices Across Days Ahead



Source: Author’s own calculations using TC-0 price data on the subset of routes described in Subsection 2.2.
Notes: The x-axis represents the number of days before departure. The y-axis presents the price of the respective ticket type expressed as a percentage of the flexible ticket price. Conditional on a given number of days ahead, this percentage is computed for each (origin, destination, calendar date of travel, departure time)-tuple. Each dot along a solid line represents the median value of this measure’s distribution. The lower and upper bounds of the shaded area represent the 25% and 75% percentiles of this measure’s distribution, respectively.

⁴A coefficient of variation is equal to the standard deviation divided by the mean.

Figure 2 shows the development of fixed ticket prices relative to flexible ticket prices over time. Generally, fixed ticket prices rise as the departure date approaches, with a notable acceleration approximately 10 to 12 days prior to departure. Most of these price increases (i.e. 98.6%) are “pre-mature” closures by the revenue manager, where price levels were closed before all allocated tickets were sold. This behaviour indicates active monitoring of available capacities and demand fluctuations by the revenue manager. It is important to note that the flexible ticket price series remains flat by design, as the analysed data subset only includes trains with constant flexible ticket prices.

An analysis of price development over days ahead, segmented by deciles of expected passenger numbers⁵ reveals an inter-temporal pattern similar to that shown in Figure 2. However, an observable difference is that starting prices are higher for trains with a greater expected number of passengers. Williams (2022) attributes these initial price level increases to rising opportunity costs. Figure A2 in Appendix A provides graphical evidence for this relationship, focusing on the first, fifth, and tenth deciles of trains by expected passenger numbers.

Overall, Figure 2, Figure A2, and the high frequency of manual price interventions motivate several key modelling assumptions. First, the observed price evolution over time and the differences in initial price levels may result from changing price sensitivities and/or increasing opportunity costs. Thus, any empirical model needs to allow for both to be present. Additionally, given the high number of manual price adjustments, I assume that only the total number of remaining discounted tickets is relevant, rather than the specific quantity at each price level. This assumption significantly reduces the computational complexity of the dynamic discrete choice model.

To further investigate the underlying drivers of pre-mature price level closures, I employ a logistic regression model to approximate the optimal policy function of the revenue manager’s dynamic pricing problem. The dependent variable is a binary indicator, equal to one if the revenue manager pre-maturely closed the respective price level, and zero otherwise. The independent variables include the lagged cumulative number of sales and cancellations. Equation 6

⁵The expected number of passengers is computed using the predicted value from a linear regression of the number of scanned tickets onto a (train origin, stops, train destination, departure time)-specific fixed effect and a (calendar date of travel)-specific fixed effect.

provides the latent variable representation for the logistic regression.

$$Y_{o,d,T,H,t} = \begin{cases} 1 & \text{if } Y_{o,d,T,H,t}^* > 0 \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

with o denoting the origin, d the destination, T the calendar date of travel, H the departure time, and t the number of days ahead. The latent variable $Y_{o,d,T,H,t}^*$ is defined below.

$$Y_{o,d,T,H,t}^* = \beta_1 \times S_{o,d,T,H,t-1} + \beta_2 \times C_{o,d,T,H,t-1} + \beta_{o,d,T} + f(\Delta_{o,d,T,H,t}) + g(p_{o,d,T,H,t-1}) + \epsilon_{o,d,T,H,t} \quad (7)$$

with $S_{o,d,T,H,t-1}$ denoting lagged cumulative sales of discounted fixed tickets, and $C_{o,d,T,H,t-1}$ denoting lagged cumulative cancellations of the same ticket type. The subscripts indicate the same as in Equation 6. $\beta_{o,d,T}$ is a (origin, destination, calendar date of travel)-specific fixed effect. The functions f and g are polynomial expansions of degree three in days ahead (i.e. $\Delta_{o,d,T,H,t}$) and lagged price (i.e. $p_{o,d,T,H,t-1}$), respectively. $\epsilon_{o,d,T,H,t}$ is an error term, unobserved to the econometrician, which is i.i.d. across all dimensions denoted by the subscripts, and follows a standard logistic distribution.

The estimation results of the logistic regression are presented in Table 2. Column 1 presents the estimation results for a logistic regression using a specification with the lagged value of remaining discounted fixed tickets. After controlling for the remaining number of days until departure (“days ahead”), and the current price level, using separate third-degree polynomial functions, and utilising a multiplicative fixed effect, one can see a statistically significant negative correlation between larger amounts of (lagged) remaining discounted fixed tickets and manual price-level closures. This negative correlation aligns with standard dynamic pricing models: as the number of remaining tickets decreases, the opportunity cost of selling an additional ticket increases, making it more likely that the revenue manager advances to a higher price level.

Column 2 decomposes the aggregate measure of remaining discounted fixed tickets into cumulative lagged sales and cumulative lagged cancellations, as specified in Equation 7. The results indicate that lagged sales are positively correlated with manual price-level closures, again consistent with standard dynamic pricing models. However, the added insight from this regression

Table 2: Motivating Evidence - Manual Fixed Ticket Price Increases

	$\mathbb{1}\{\text{Manual Price-Level Closure}\}$	
	(1)	(2)
Remaining Tickets (Lagged)	-0.004*** (1.9E-04)	-
Cumulative Ticket Sales (Lagged)	-	0.056*** (0.002)
Cumulative Ticket Cancellations (Lagged)	-	-0.010*** (0.002)
Poly. Degree 3 - Days Ahead	✓	✓
Poly. Degree 3 - Price (Lagged)	✓	✓
Fixed Effect	Origin × Destination × Calendar Date of Travel	Origin × Destination × Calendar Date of Travel

Source: Author's own calculations using TC-0 price, sales, and cancellation data for the subset described in Subsection 2.2. Only discounted fixed tickets are used.

Notes: An observation is defined at the (origin, destination, calendar date of travel, departure time, day ahead)-level. Standard errors are clustered at the (origin, destination, calendar date of travel)-level, and provided in parentheses. The dependent variable in both Columns 1 and 2 is equal to one if the respective fixed ticket price level was closed pre-maturely, and zero elsewhere. Column 1 features the lagged number of remaining discounted fixed tickets as an explanatory variable. Column 2 features the lagged number of cumulative discounted fixed ticket sales and the lagged number of cumulative discounted fixed ticket cancellations as explanatory variables. The fixed effect across both specifications is the interaction between the origin, destination, and calendar date of travel.

* ... $p < 0.05$ ** ... $p < 0.01$ *** ... $p < 0.001$

is the statistically significant negative correlation between lagged cancellations and price-level closures. As the revenue manager regains capacity through cancellations, the opportunity cost of selling seats decreases, reducing the likelihood of price increases. Cancellations thus act as a countervailing force on the opportunity cost of selling goods with limited capacity. A structural model of dynamic pricing must therefore incorporate this channel to accurately capture the interplay of cancellations and pricing decisions.

4 Empirical Model

In this section, I develop an empirically estimatable structural model of dynamic train-level pricing that accounts for cancellations. Subsection 4.1 outlines the general framework of the model. Subsection 4.2 provides a detailed analysis of the demand-side components, while Subsection 4.3 focuses on the supply-side aspects of the structural model.

4.1 General Setting

The structural model developed here focuses on dynamic pricing at the train-level. Specifically, it applies to a single train departing at time H for non-stop travel between an origin station o and a destination station d on a specified calendar date of travel T . On any day prior to departure, denoted by Δ , consumers face a choice between two distinct products: a fixed ticket for the designated train, or a flexible ticket valid for travel between the two stations on the specified calendar date of travel. The revenue manager of the passenger rail company actively determines the price for a pre-specified quantity of discounted fixed tickets, while the price of the flexible ticket is treated as exogenously given and constant throughout the sales period. For clarity and to enhance readability, subscripts denoting the origin (o), destination (d), calendar date of travel (T), and departure time (H) are omitted in the remainder of this section.

The model begins 28 days prior to the train's departure, i.e. $\bar{\Delta} = 28$, and progresses through each day until the day of departure, i.e. $\Delta = 0$. Each time period corresponds to a full day. The timing of events within each day is illustrated in Figure 3. At the start of each day, before consumers make their purchasing decisions, the revenue manager observes sales and cancellations from the previous period. Based on this information, the manager decides whether to maintain the current price level for fixed tickets or to move to the next higher price level. Following this pricing decision, consumers arrive and choose between purchasing a fixed ticket, a flexible ticket, or no train ticket at all. After these purchasing decisions, all ticket holders are subject to the possibility of an adverse "cancellation event". In such cases, affected consumers must cancel their tickets immediately.

Figure 3: Timing Within One Period



4.2 Demand

Arrival Each day prior to departure, a stochastic number of consumers visits the TC-0 website. The number of daily arrivals is modelled as a Poisson random variable, with the mean arrival rate varying based on the number of days remaining until departure.

$$\mathbf{A}_\Delta \sim Pois(\lambda_\Delta) \quad (8)$$

Indirect Utility Upon arrival, each consumer i faces a discrete choice among three options: purchasing a fixed ticket, purchasing a flexible ticket, or not purchasing a ticket. The consumer's choice is modelled using a discrete choice random utility framework, where the indirect utility is linear in product characteristics and idiosyncratic preferences. The indirect utility function for both the fixed and flexible tickets is specified in Equation 9, while the indirect utility of not buying is normalised to $u_{i,\Delta,0} = 0 + \epsilon_{i,0}$.

$$\begin{aligned} u_{i,\Delta,j} &= EU_{\Delta,j} + \epsilon_{i,j} \\ &= \alpha_\Delta \times -p_{\Delta,j} + \mu_\Delta \times (\delta_j) + (1 - \mu_\Delta) \times (\alpha_\Delta \times (-f_{\Delta,j} + p_{\Delta,j})) + \epsilon_{i,j} \end{aligned} \quad (9)$$

with j denoting the ticket type.

Equation 9 decomposes the indirect utility into three main components, in addition to an idiosyncratic preference shock. The first component is the immediate impact on indirect utility from paying the ticket price $p_{j,\Delta}$. This represents the disutility of expenditure and is scaled by the consumer's price sensitivity. The second component is the "benefit of travel". This term captures the indirect utility of travelling, net of price effects. Benefits of travel may vary by ticket type and are weighted by the probability of travelling, μ_Δ . The last component is the effect on indirect utility in the event of cancellation. This component accounts for the utility

consequences in the event of a ticket cancellation. Upon cancellation, the consumer pays a cancellation fee $f_{j,\Delta}$ but receives a refund of their original payment $p_{j,\Delta}$. The refunded amount is scaled by the price sensitivity parameter α_Δ , reflecting the perception of receiving money back. The probability of cancellation, or not travelling, is $(1 - \mu_\Delta)$.

Discounting, and Risk Neutrality Given the relatively short time horizon of at most 28 days prior to departure, I abstract away from time-discounting both the benefits of travel and the reimbursement amount. The rationale is that, within such a brief period, the perceived utility of future events and the time value of money are unlikely to be meaningfully discounted by consumers. These outcomes are, however, already implicitly discounted through the increased uncertainty associated with booking early, thus negating the need for explicit time-discounting in the model.

The linear specification of the indirect utility function implicitly assumes risk neutrality. This assumption is reasonable in this context, as the monetary stakes involved - specifically, at most the cancellation fee $f_{j,\Delta}$ - are relatively small. Given the limited financial risk, it is plausible to treat the (indirect) utility function as approximately linear over the relevant range of potential wealth loss. Consequently, deviations from risk neutrality are unlikely to meaningfully affect consumer decision-making in this setting.

Choice An individual consumer is assumed to select the option that provides the highest indirect utility to her. Since the vector of idiosyncratic preference shocks, ϵ , is unobservable to both the revenue manager and the econometrician, I assume that the components of ϵ are independently and identically distributed (i.i.d.) across consumers and choices, following a Type 1 Extreme Value (T1EV) distribution. The probability that consumer i chooses ticket option j , conditional on Δ days ahead, is thus given by Equation 10.

$$\mathbb{P}(i \text{ chooses } j|\Delta) = \frac{\exp(EU_{\Delta,j})}{1 + \exp(EU_{\Delta,fix}) + \exp(EU_{\Delta,flex})} \quad (10)$$

4.3 Dynamic Pricing

The revenue manager at Train Company Zero (TC-0) is assumed to maximise the expected revenue for a given train, subject to a quantity constraint on discounted fixed tickets. Specifically, a finite quantity of tickets, \bar{N} , is available for sale at prices below the "terminal" price level, \bar{p}_{fix} . The manager's decision-making process follows a dynamic optimisation framework. She adjusts ticket prices according to a predetermined price schedule, \mathcal{P} , where prices can only increase; downward price revisions are precluded by the system's constraints. At each decision point, the manager must make a forward-looking choice: either maintain the current price level or advance to the next (higher) price in the schedule. The price of flexible tickets, \bar{p}_{flex} , is assumed to remain constant throughout the sale period, allowing the manager to focus solely on optimising the pricing trajectory for fixed tickets. Furthermore, it is assumed that the revenue manager possesses complete knowledge of all relevant demand parameters, including their evolution over time.

As described in Subsection 4.1, the revenue manager's decision-making process is guided by observed state variables, which are updated daily based on the prior day's sales and cancellations. The state variables that influence her decisions on any given day are: the number of still available discounted tickets N , the number of days until departure Δ , the current fixed ticket price level p_{fix} , and a set of choice-specific revenue shocks $\tilde{\epsilon}$. After making a pricing decision, the current day's events unfold: consumers arrive, sales and cancellations are realised, and the state variables are updated accordingly. This decision-making repeats each day until one of three terminal conditions is met: the day of departure arrives ($\Delta = 0$), the highest price level is reached ($p_{fix} = \bar{p}_{fix}$), or the supply of discounted tickets is exhausted ($N = 0$). Upon reaching any of these terminal conditions, the pricing defaults to the predetermined levels ($\bar{p}_{fix}, \bar{p}_{flex}$), and the revenue manager ceases active decision-making for the train in question.

The recursive structure of this pricing problem, coupled with the binary nature of the daily pricing decision, allows the problem to be formulated as a dynamic discrete choice (DDC) problem. The corresponding value function, described by Equation 11, captures the manager's optimisation framework and represents the maximum expected revenue attainable from the

current state onward.

$$V(N, \Delta, p_{fix}, \tilde{\epsilon}) = \max \begin{cases} \mathbb{E}[R(p_{fix}, N, \Delta)] + \tilde{\epsilon}_{p_{fix}} + \beta \times \mathbb{E}[V(N', \Delta', p_{fix}, \tilde{\epsilon}')] & \text{maintain} \\ \mathbb{E}[R(\hat{p}_{fix}, N, \Delta)] + \tilde{\epsilon}_{\hat{p}_{fix}} + \beta \times \mathbb{E}[V(N', \Delta', \hat{p}_{fix}, \tilde{\epsilon}')] & \text{higher} \end{cases} \quad (11)$$

with N denoting the remaining number of discounted fixed tickets, Δ denoting the remaining days until departure, p_{fix} denoting the current fixed ticket price level, and $\tilde{\epsilon}$ denoting the vector of action-specific revenue shocks only observable to the revenue manager. β represents the discount factor, and thus governs the trade-off between current and future revenue.

$\mathbb{E}[R(p_{fix}, N, \Delta)]$ denotes the instantaneous expected revenue function. In addition to the time-invariant flexible ticket price \bar{p}_{flex} , instantaneous revenue also depends on the time-varying parameters of the demand function as well as the current fixed ticket price p_{fix} and the number of discounted tickets N still available. The specification of $\mathbb{E}[R(p_{fix}, N, \Delta)]$ is given by Equation 12.

$$\begin{aligned} \mathbb{E}[R(p_{fix}, N, \Delta)] = & \mathbb{E}[Q_{fix}|Q_{fix} \leq N] * (p_{fix} + (1 - \mu) \times (f_{fix} - p_{fix})) + \\ & \mathbb{E}[Q_{fix}|Q_{fix} > N] * (\bar{p}_{fix} + (1 - \mu) \times (f_{fix} - \bar{p}_{fix})) + \\ & \mathbb{E}[Q_{flex}] * (\bar{p}_{flex} + (1 - \mu) \times (f_{flex} - \bar{p}_{flex})) \end{aligned} \quad (12)$$

The first term in Equation 12 represents the revenue derived from discounted fixed ticket sales at the current price level p_{fix} , adjusted for the expected cancellation fees associated with these tickets. The second term accounts for revenue from fixed ticket sales, also net of expected cancellation fees, but in the event that the supply of discounted fixed tickets is exhausted during the same period. In such cases, the fixed ticket price defaults to \bar{p}_{fix} , as described previously. The final term captures revenue from flexible ticket sales, net of expected cancellation fees.

The “technical” marginal cost of transporting another passenger is normalised to zero, simplifying the cost structure to focus exclusively on the opportunity cost of selling the seat. This opportunity cost reflects the foregone revenue from potentially selling the (discounted) fixed ticket in a future period.

The transition equation for the number of available discounted fixed tickets, N , is given below by Equation 13. Contrary to a standard capacity transition equation, Equation 13 allows for capacity to be recovered through cancellations. Specifically, if the realised number of discounted fixed ticket cancellations, C_{fix} , is larger than the realised number of discounted fixed ticket sales, Q_{fix} , the number of available tickets in the next period, N' , may increase relative to the current period's, N . However, the support of N' is constrained to $\{0, 1, \dots, \bar{N}\}$, with \bar{N} denoting the total number of discounted fixed tickets allocated to the train.

$$N' = N + C_{fix} - Q_{fix} \quad (13)$$

The number of days remaining until departure, Δ , evolves deterministically according to Equation 14. The support is restricted to $\{0, 1, \dots, \bar{\Delta} = 28\}$.

$$\Delta' = \Delta - 1 \quad (14)$$

The terminal conditions described previously are formalised as follows. The first terminal condition describes the circumstance that on the day of departure, all remaining fixed tickets - whether discounted or not - are sold at the terminal price \bar{p}_{fix} . The second terminal condition captures the circumstance that if the revenue manager chooses to advance the fixed ticket price to the highest price level \bar{p}_{fix} before the day of departure, all fixed tickets are sold at \bar{p}_{fix} until the train departs. The third terminal condition states that once the inventory of discounted tickets is depleted, the pricing system defaults to the terminal prices $(\bar{p}_{fix}, \bar{p}_{flex})$ for all subsequent periods. Please note the introduction of a new capacity variable, i.e. $\bar{\bar{N}}$, which denotes the

technical capacity of the train.

$$\begin{aligned}
V(N, 0, p, \epsilon) &= \mathbb{E}[R(\bar{p}_{fix}, \bar{N} - N, 0)] \\
V(N, \Delta, \bar{p}_{fix}, \epsilon) &= \sum_{k=0}^{\Delta} \beta^k \times \mathbb{E}[R(\bar{p}_{fix}, \bar{N} - N, \Delta - k)] \\
V(0, \Delta, p_{fix}, \epsilon) &= \sum_{k=0}^{\Delta} \beta^k \times \mathbb{E}[R(\bar{p}_{fix}, \bar{N} - \bar{N}, \Delta - k)]
\end{aligned}$$

Conditional Choice Probability To derive an analytical form for both the expected value function tomorrow (conditional on today's state variables and action) and the choice probabilities for today's decisions, I impose several assumptions. First, I make the conditional independence assumption following Rust (1987). This assumption implies that the transition probabilities of the unobserved revenue shock state variables, ϵ , are conditionally independent of the transition probabilities of the observed state variables. Second, I assume that the choice-specific revenue shocks, ϵ_p and $\epsilon_{\hat{p}}$, which are unobservable to the econometrician, are independently and identically distributed (i.i.d.) across choices and state-variable combinations. These shocks are further assumed to follow a Type-1 Extreme Value (T1EV) distribution.

Given the assumptions, the expected value function conditional on today's state variables is given by Equation 15.

$$\begin{aligned}
EV(N, \Delta, p_{fix}) &= \sum_{N'} \left[\sigma \times \log \left(\sum_{p' \in \{p_{fix}, \hat{p}_{fix}\}} \exp \left(\frac{\mathbb{E}[R(N', p', \Delta')] + \beta \times EV(N', \Delta', p')}{\sigma} \right) \right) \right] \\
&\quad \times g(N', \Delta' | N, \Delta, p) + \sigma \times \gamma
\end{aligned} \tag{15}$$

with σ being the scale factor of the revenue shocks' T1EV distribution, $g(\cdot)$ denoting the probability mass function (PMF) of capacity transitions, and γ being Euler's constant.

The probability of increasing the fixed ticket's price in the current period, given the current

vector of observable state variables, is given by Equation 16.

$$\mathbb{P}(\hat{p}_{fix}|N, \Delta, p_{fix}) = \frac{\exp\left(\left(\mathbb{E}[R(\hat{p}_{fix}, N)]\right) + \beta \times EV(N, \Delta, \hat{p})\right)/\sigma}{\sum_{p' \in \{p_{fix}, \hat{p}_{fix}\}} \exp\left(\left(\mathbb{E}[R(p', N)]\right) + \beta \times EV(N, \Delta, p')\right)/\sigma)} \quad (16)$$

5 Estimation

This section outlines the empirical parameterisation and estimation strategy for the structural model of demand and supply developed in Section 4. The model parameters are estimated separately for each (o,d,w)-tuple, where o denotes the origin, d the destination, and w the weekday of travel. To streamline the exposition, I omit these subscripts in the text below. Within each (o,d,w)-tuple, variation across departure times and calendar days is utilised to identify and estimate the model parameters.

5.1 Cancellation Risk

The probability of travel, μ_Δ , is conceptualised as the outcome of a series of repeated Bernoulli trials. Each day, a ticket holder receives a signal generated by an exogenous process. Upon receiving an adverse cancellation signal, the ticket holder is compelled to cancel her ticket immediately. The daily cancellation risk is characterised by the parameter τ , which may vary across days leading up to the departure date. In accordance with the cancellation policy of TC-0 and the structural model outlined above, cancellation is not possible on the day of travel. Consequently, a ticket holder who remains unaffected by cancellation signals as of the final pre-travel day is assumed to travel the following day⁶. Assuming the Bernoulli trials are independent across days, the probability of travel is derived in Equation 17.

$$\mu_\Delta = \prod_{z=\Delta}^1 (1 - \tau_z) \quad (17)$$

The daily cancellation risk, τ_z , is further parameterised using the standard logistic function. This functional form ensures that τ_z is confined to the interval $[0, 1]$, while allowing for sufficient

⁶In the dataset, less than 0.25% of non-cancelled tickets are never validated. These observations are treated as negligible in the estimation process.

flexibility for the daily risk to vary over time.

$$\tau_z = \frac{\exp(\tilde{\tau}_1 + \tilde{\tau}_2 \times z + \tilde{\tau}_3 \times z^2)}{1 + \exp(\tilde{\tau}_1 + \tilde{\tau}_2 \times z + \tilde{\tau}_3 \times z^2)} \quad (18)$$

Estimation The parameters governing the daily risk of cancellation, i.e. $\tilde{\tau}$, are estimated using maximum likelihood estimation (MLE) with the probability mass function (PMF) of the geometric distribution. The likelihood contribution l_i of individual i is given by Equation 19.

$$l_i(\tilde{\tau}) = \begin{cases} \prod_{z=\Delta_i}^{X-1} (1 - \tau_z(\tilde{\tau})) \times \tau_{\Delta_i-X}(\tilde{\tau}) & \text{consumer } i \text{ cancelled after } X \text{ days} \\ \prod_{z=\Delta_i}^1 (1 - \tau_z(\tilde{\tau})) & \text{consumer } i \text{ never cancels} \end{cases} \quad (19)$$

Identification The parameter vector $\tilde{\tau}$ is identified through variation in observed cancellation rates across different departure times and calendar days within the relevant (origin o , destination d , weekday of travel w)-tuple. Leveraging precise timestamps for both the time of purchase and potential cancellation, as well as the known cancellation deadline for each train, the exact duration for which each ticket was held can be accurately recovered.

5.2 Demand Parameters

The time-varying mean of the Poisson arrival process is parameterised as specified in Equation 20. To ensure the non-negativity of the Poisson distribution's mean, the exponential function is employed. Similar to the approach outlined above, a quadratic specification is imposed on the mean arrival rate to allow for flexible variation over time.

$$\lambda_{\Delta} = \exp(\tilde{\lambda}_1 + \tilde{\lambda}_2 \times \Delta + \tilde{\lambda}_3 \times \Delta^2) \quad (20)$$

The mean utilities of train tickets net of price, referred to as the "benefits of travel," are defined in Equation 21. The parameter δ_1 is interpreted as the baseline benefit of travel, representing the common utility component shared by both ticket types. The primary distinction between a fixed and a flexible ticket, aside from price or cancellation fees, lies in the freedom to select a departure time. Accordingly, δ_2 can be interpreted as the incremental utility derived from the

flexibility to choose one's departure time.

$$\begin{aligned}\delta_{fixed} &= \delta_1 \\ \delta_{flexible} &= \delta_1 + \delta_2\end{aligned}\tag{21}$$

The price sensitivity parameter is specified to vary across days prior to departure, as outlined in Equation 22.

$$\alpha_\Delta = \tilde{\alpha}_1 + \tilde{\alpha}_2 \times \Delta + \tilde{\alpha}_3 \times \Delta^2\tag{22}$$

Estimation The estimation of demand parameters follows a two-step procedure. Due to the presence of two distinct product types for each train - fixed and flexible tickets - and the assumed logit demand system, several demand parameters can be estimated independently of the full dynamic discrete choice (DDC) model. Specifically, the price sensitivity parameters, $\tilde{\alpha}$, and the utility benefit of flexible departure times, δ_2 , are estimated using maximum likelihood estimation (MLE) using individual purchase data. The individual choice probabilities required for the computation of the likelihood function are defined in Equation 10 in Subsection 4.2.

These first-stage estimates are utilised to ease the computational burden of estimating the full DDC model. The remaining demand parameters, δ_1 and $\tilde{\lambda}$, are estimated using a modified version of the nested fixed-point (NFXP) algorithm introduced by Rust (1987). Given the finite horizon of the dynamic pricing problem, the value function for a given parameter vector is solved through backward induction, as opposed to fixed-point iteration. For a given parameter vector, the model provides a closed-form expression for the conditional choice probabilities (CCPs), as specified in Equation 16. These CCPs enable the computation of the likelihood function, allowing the remaining parameters to be estimated via MLE.

The likelihood is given below by Equation 23.

$$l(\delta_1, \tilde{\lambda}) = \prod_{tr \in \mathcal{TR}} \prod_{\Delta} \left[\left(\mathbb{P}(a_{tr,\Delta} | N_{tr,\Delta}, \Delta, p_{fix,tr,\Delta}) \right) \times \left(g(N_{tr,\Delta-1} | N_{tr,\Delta}, \Delta, p_{fix,tr,\Delta}, a_{tr,\Delta}) \right) \right]\tag{23}$$

with \mathcal{TR} denoting the set of all trains within the relevant (origin o , destination d , weekday of

travel w)-tuple.

Identification The parameters δ_2 and $\tilde{\alpha}$ are identified from consumers' choices between fixed and flexible tickets. Given that the probability of travel is derived using the estimates from Subsection 6.1 and the cancellation fee structure is observable, consumers face a trade-off between the additional benefit of flexibility, δ_2 , and the higher price associated with flexible tickets. Following the principle that only differences in utilities are relevant in discrete choice models, and leveraging the independence of irrelevant alternatives (IIA) property inherent to logit demand systems, δ_2 and $\tilde{\alpha}$ are identified from consumers' decisions between the two ticket types⁷. However, δ_1 remains unidentified when only purchase data are used. As noted by Williams (2022), this limitation arises from the absence of "no-purchase" data, which precludes the estimation of the baseline utility of travel.

I build on the identification argument developed by Williams (2022) to identify δ_1 and the parameters governing the potential market size, $\tilde{\lambda}$. First, I assume that TC-0 revenue managers possess full knowledge of how demand parameters evolve over time, and that their observed pricing behaviour is optimal given this information. Second, within the relevant (origin o , destination d , weekday of travel w)-tuple - conditional on the days prior to departure - and having quantified the opportunity cost of selling a seat (i.e., the marginal cost) by solving the model, the remaining variation in the number of discounted tickets available, N , and the fixed ticket price level, p_{fix} , as well as the subsequent pricing decisions by the revenue manager, provides information about the unobserved demand parameters.

The revenue manager's discount factor, β , is assumed to be 0.999. Given that the dynamic discrete choice model is finite-horizon, it is solved via backward induction and does not rely on the strict $\beta < 1$ assumption required for contraction mapping. The horizon for decision-making is set at a maximum of 28 days prior to departure, as described earlier. Furthermore, since TC-0 does not engage in strategic overbooking of trains, the model excludes this consideration.

⁷This identification strategy builds on the framework established by Berry (1994). Rather than differencing the logarithm of a ticket's market share relative to the logarithm of the outside option's market share, the logarithm of the flexible ticket's market share is differenced relative to that of the fixed ticket. This allows for the recovery of the difference in mean utilities, thereby identifying δ_2 and $\tilde{\alpha}$, though not δ_1 .

Unobserved Heterogeneity The demand model does not incorporate unobserved heterogeneity in any of the structural parameters. This exclusion is motivated by the substantial computational complexity that would arise from allowing for individual-specific travel probabilities and/or price sensitivities. Accounting for individual-specific travel probabilities would necessitate tracking all previously purchased tickets as state variables within the dynamic discrete choice (DDC) model to compute the cancellation probabilities for the current period. This would lead to a significant expansion in the dimensionality of the state space. Similarly, introducing unobserved heterogeneity in the price sensitivity parameter would undermine the first-stage estimation of the price sensitivity parameters, requiring all parameters to be estimated jointly within the DDC framework and further compounding computational complexity. That said, the model already allows for considerable flexibility: the parameters vary at the (origin o , destination d , weekday of travel w)-level, while price sensitivities and cancellation risks additionally vary across days prior to departure.

Consumer Choice Set As outlined previously, the model restricts consumer choices to either a fixed ticket for a specific train or a flexible ticket valid for the calendar date of travel. Consequently, the possibility of substitution between different fixed tickets (i.e. departure times) on the same day is excluded. Expanding the choice set to include multiple departure times would require that the revenue manager accounts for these substitution patterns when making pricing decisions, thereby transforming the problem into a joint dynamic pricing problem across multiple departure times. Such an extension would substantially increase the dimensionality of the state space, as it would require tracking the availability of discounted fixed tickets for all departure times. Moreover, the action space would expand significantly, as all potential combinations of pricing decisions across the departure schedule would need to be considered.

6 Empirical Results

This section presents the estimation results for the empirical model outlined in Section 4. The model is estimated across all (origin, destination, weekday of travel)-tuples. For brevity, results are reported for a subset of four selected tuples. These tuples include two weekday and two

weekend travel days, chosen to provide diversity in terms of travel times.

Due to a confidentiality clause in the data-sharing agreement, the precise number of observations in each tuple cannot be disclosed, as this information could potentially allow for the identification of TC-0. However, it can be confirmed that each of the four selected tuples comprises more than one hundred thousand observations, providing a robust basis for estimation.

6.1 Cancellation Risk Estimates

Table 3: Empirical Results - Daily Cancellation Risk

	(1)	(2)	(3)	(4)
Risk of Cancellation				
$\tilde{\gamma}_1$	-4.448*** (0.079)	-4.010*** (0.076)	-4.192*** (0.096)	-4.146*** (0.092)
$\tilde{\gamma}_2$	-0.172*** (0.024)	-0.194*** (0.025)	-0.142*** (0.032)	-0.262*** (0.027)
$\tilde{\gamma}_3$	0.004*** (0.001)	0.004** (0.001)	0.002 (0.002)	0.009*** (0.001)
Route	A to B	C to A	A to C	B to A
Weekday	Friday	Saturday	Wednesday	Tuesday

Source: Author's own calculations using TC-0 cancellation data.

Notes: Standard errors in parentheses.

* ... $p < 0.05$ ** ... $p < 0.01$ *** ... $p < 0.001$

Table 3 reports the daily cancellation risk parameters for the four selected (origin, destination, weekday of travel)-tuples. Within each tuple, a clear pattern emerges: daily cancellation risks are lower for bookings made far in advance of the departure date compared to those made closer to departure. However, due to the compounding nature of individual cancellation risks, the actual travel probability for early bookers is naturally lower than that for late bookers. A visualisation of the travel probability, μ_Δ , by days ahead and tuple is provided in Figure A3 in Appendix A.

Across the selected tuples, there is notable heterogeneity in the daily cancellation risk parameters. Comparing the tuples 28 days prior to departure reveals significant differences in the overall probability of travel. These disparities persist, albeit to a diminishing extent, until

approximately seven days before departure, at which point travel probabilities begin to converge. On the final day before departure - the last day when cancellations are permitted - these differences become less pronounced but remain observable.

Selection As outlined in Subsection 5.1, the model does not account for selection into ticket types driven by unobserved heterogeneity in cancellation risk, owing to the associated computational complexity. Consumers with a high likelihood of cancellation may prefer flexible tickets, as these allow greater flexibility in departure time and offer full reimbursement (i.e. a zero cancellation fee). To address this concern, I estimate the cancellation risk parameters using only fixed-ticket holders. The results are presented in Table A1 in Appendix A. Across columns, the parameter estimates are similar to those obtained from the full sample of cancellations. While this does not constitute a definitive test for selection into ticket types, it provides empirical evidence suggesting no substantial differences in observed cancellation behaviour.

Another potential source of consumer selection arises from the actual cancellation decision. Consumers who paid a higher price may have greater financial incentives to cancel and recover their costs upon experiencing an adverse cancellation event. However, this is unlikely, as cancellation requires no justification and can be completed with minimal effort. To test this hypothesis empirically, I analyse all tickets sold prior to the cancellation deadline but never used or validated. This dataset includes both cancelled (refunded) tickets and tickets eligible for cancellation but not cancelled. Regressing a cancellation indicator on the ticket price and a fixed effect for (origin, destination, weekday of travel, days ahead)-specific fixed effect, I find no statistically significant relationship between the original ticket price and the likelihood of cancellation. These results are presented in Table A2 in Appendix A, which reports estimates from both a linear probability model and a logistic regression.

6.2 Demand Parameter Estimates

The demand estimates for the four selected (origin, destination, weekday of travel)-tuples are reported in Table 4. The first panel presents the parameter estimates governing the time-varying mean of the consumer arrival process. The second panel provides the estimates for the

Table 4: Empirical Results - Demand

	(1)	(2)	(3)	(4)
Arrival Rate				
$\tilde{\lambda}_1$	2.358*** (0.022)	2.089*** (0.023)	3.007*** (0.036)	2.251*** (0.023)
$\tilde{\lambda}_2$	-0.407*** (0.007)	-0.516*** (0.008)	-0.373*** (0.009)	-0.304*** (0.008)
$\tilde{\lambda}_3$	0.010*** (2.9E-04)	0.016*** (3.5E-04)	0.010*** (4.3E-04)	0.006*** (3.3E-04)
Price Sensitivity				
$\tilde{\alpha}_1$	-0.032*** (0.005)	-0.111*** (0.007)	-0.039*** (0.006)	-0.044*** (0.005)
$\tilde{\alpha}_2$	-0.009*** (0.001)	-0.007*** (0.001)	-0.006*** (0.001)	-0.005*** (0.001)
$\tilde{\alpha}_3$	3.0E-04*** (2.4E-05)	9.2E-05 (5.4E-05)	8.1E-05* (3.9E-05)	1.6E-04*** (2.6E-05)
Benefits of Travel				
δ_1	2.153*** (0.045)	3.068*** (0.018)	-1.406*** (0.042)	1.083*** (0.017)
δ_2	2.166*** (0.038)	2.102*** (0.045)	2.184*** (0.051)	2.275*** (0.047)
Route	A to B	C to A	A to C	B to A
Weekday	Friday	Saturday	Wednesday	Tuesday

Source: Author's own calculations using TC-0 sales, pricing, and cancellation data.

Notes: Standard errors in parentheses.

* ... $p < 0.05$ ** ... $p < 0.01$ *** ... $p < 0.001$

time-varying price sensitivity parameters. Finally, the third panel reports the estimates for the "benefits of travel."

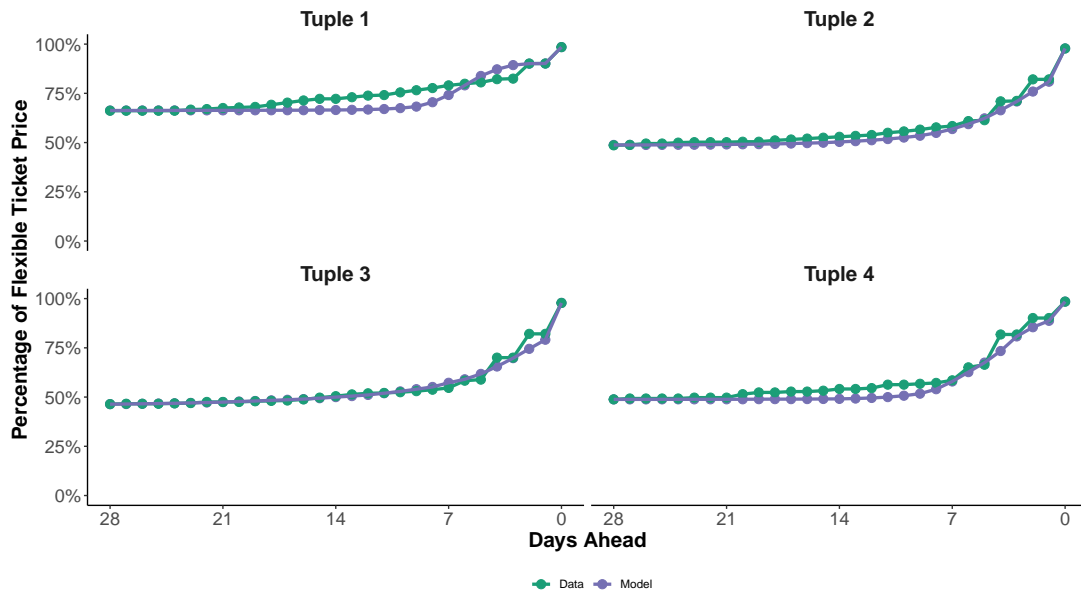
The mean arrival rates of potential consumers exhibit significant variation across days prior to departure. While there is some heterogeneity in arrival rates 28 days in advance, the majority of between-tuple differences emerge within the final seven days before departure. Overall, the results indicate that most potential consumers arrive closer to the train's departure date, with the highest mean arrival rate observed on the day of departure. Figure A4 in Appendix A visualises the mean arrival rates over days ahead for each selected tuple. Notably, the figure

also suggests that between 21 and 14 days prior to departure, almost no potential consumers arrive at TC-0's website. In contrast, Williams (2022) finds that in the U.S. airline industry, the number of potential consumers arriving early exceeds those arriving closer to the departure deadline. This discrepancy highlights notable differences between consumer arrival processes in the rail and aviation industries.

The middle panel of Table 4 illustrates the evolution of price sensitivities across days ahead. Similar to the mean arrival rates, there is substantial heterogeneity in price sensitivity over time. The tuples represented in Columns (1) and (4) generally exhibit less price-sensitive consumers at both the beginning and the end of the booking period. In contrast, the price sensitivity of consumers in the tuples represented in Columns (2) and (3) evolves differently: the sensitivity in Column (3) declines as the departure date approaches, while the tuple in Column (2) is characterised by consistently higher levels of price sensitivity throughout the entire pre-departure period. In comparison to Column (2), the magnitude of price sensitivity for the other three tuples is more aligned. A visual representation of the price sensitivity parameter across days ahead for each tuple is provided in Figure A5 in Appendix A.

The bottom panel of Table 4 presents the estimates for the benefits of travel. With the exception of Column (3), the common benefit of travel shared by both ticket types is positive, indicating that consumers would most likely travel if tickets were free. In the case of Column (3), however, the common benefit of travel is negative. Since these estimates are relative to an outside option normalised to a mean utility of zero, this suggests that even with a free fixed ticket, the outside option is strongly preferred. The magnitude of δ_2 , which represents the additional utility from flexible departure times, is comparable across all four tuples. This similarity is likely attributable to the consistent service frequency across the selected (origin, destination, weekday of travel)-tuples⁸.

Figure 4: Empirical Pricing Behaviour vs. Model-Predicted Pricing Behaviour



Source: Author’s own illustration using TC-0 price data as well as simulation outcomes.

Notes: The x-axis represents the days before departure. The y-axis represents the fixed ticket price as a percentage of the time-constant flexible ticket price. Each panel represents a different (origin, destination, weekday of travel)-tuple. Within each panel, the green line is the empirical average fixed ticket price, while purple is the model-simulated average fixed ticket price.

6.3 Model Fit

Figure 4 compares observed pricing behaviour in the data with simulated pricing behaviour generated using the structural dynamic pricing model and the parameter estimates discussed above. For each of the four (origin, destination, weekday of travel)-tuple, behaviour is simulated separately based on 1,000 replications. Overall, the model captures the observed pricing behaviour of TC-0 revenue managers effectively. Across all four tuples, the mean simulated price paths align closely with their empirical counterparts, demonstrating the model’s ability to replicate key patterns in the data.

⁸While not included in the main results, a regression of δ_2 on the number of available services yields a statistically significant positive coefficient. This indicates that the utility derived from flexibility increases with the number of departure options.

7 Counterfactuals

The objective of this paper is to empirically assess the impact of the flexibility to cancel on ticket prices, consumer welfare, and firm revenue within a dynamic pricing framework. Utilising the structural dynamic pricing model developed in Section 4 and Section 5, and estimated in Section 6, I conduct a series of counterfactual scenarios to analyse these effects.

Currently, TC-0 permits all ticket holders to cancel their tickets up to one day before the scheduled departure. In the baseline scenario for any of the four tuples, both consumer-side cancellation and firm-side re-selling are allowed. This baseline scenario corresponds to Equilibrium II in Figure 5. In contrast, Equilibrium I represents a counterfactual scenario in which neither cancellations nor re-selling are permitted. Computing optimal prices, consumer welfare, and firm revenue under this counterfactual allows for an empirical quantification of the value of flexibility. Additionally, two intermediate counterfactual scenarios - Equilibria III and IV in Figure 5 - are employed to decompose the changes from Equilibrium II to Equilibrium I. This decomposition provides insights into the relative contributions of consumer-side cancellations and firm-side re-selling to overall outcomes.

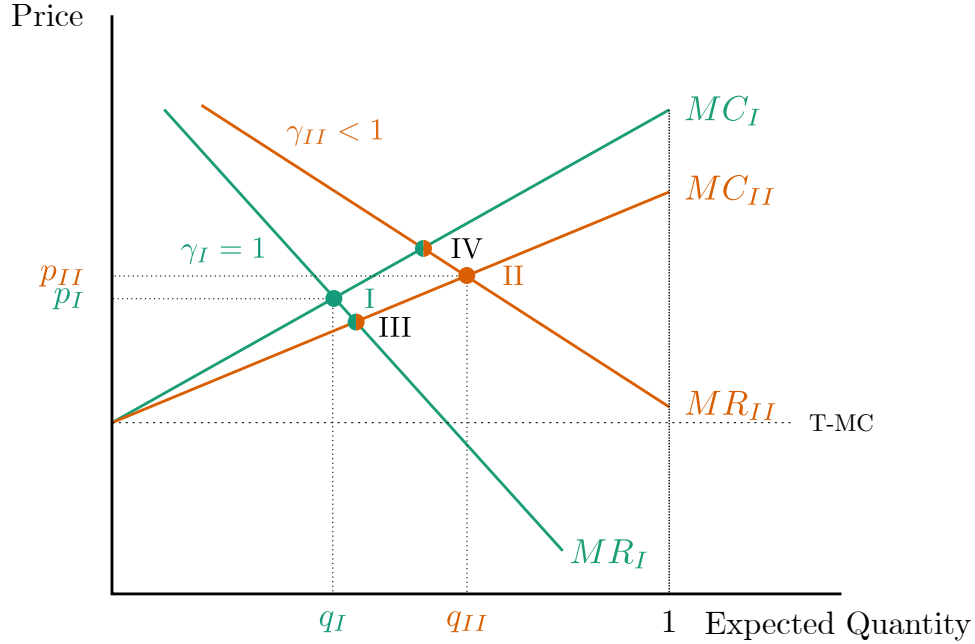
Equilibrium III allows TC-0 to re-sell cancelled tickets but does not reimburse consumers in the event of cancellation. Under this setting, opportunity costs remain constant relative to the baseline scenario of Equilibrium II. Consequently, any changes observed arise from the rotation of the marginal revenue curve, a mechanism I refer to as the ticket quality channel.

Equilibrium IV, by contrast, focuses on the role of opportunity costs. In this scenario, TC-0 reimburses consumers for cancellations but is prohibited from re-selling cancelled tickets. Any deviations from Equilibrium II in this case are driven entirely by changes in opportunity costs, a mechanism I refer to as the opportunity cost channel.

By analysing these two intermediate counterfactuals relative to the baseline, I decompose the net effect into the contributions of the ticket quality channel and the opportunity cost channel.

Technical Implementation Using the estimated parameters from Section 6, I simulate the model 1,000 times for each tuple, holding the random draws constant across all counterfactual

Figure 5: Decomposition Counterfactuals



Source: Authors' own illustrations.

Notes: Subscript I indicates the situation without cancellation, while Subscript II indicates a situation with cancellation. The marginal revenue curves are labelled MR , and the marginal cost curves are labelled MC . The horizontal line labelled $T-MC$ denotes the technical marginal cost of providing the ticket. The variable γ indicates the share of the original price that the consumer must pay if she cancels her ticket, i.e. the cancellation fee. Please note that I have not explicitly drawn the discontinuity in MC_I , MC_{II} , $T - MC$ at $\mathbb{E}[q] = 0$.

scenarios. Similar to the empirical setting, simulations commence 28 days prior to departure. The initial total number of discounted fixed tickets, their respective pricing menus and starting prices, as well as the flexible ticket prices, are drawn from the tuple-specific joint empirical distribution.

Within each tuple and simulation, I use fixed ticket prices, consumer welfare, and firm revenue from the baseline scenario (i.e. Equilibrium II) as benchmark metrics. Holding random draws constant, I then vary the model structure according to each counterfactual scenario to generate new prices, consumer welfare, and firm revenue. Deviations from the benchmark are computed for all days except the day of departure (days ahead equal to zero). Observations for the day of departure are excluded because cancellations are no longer possible, and prices are fixed under the implemented pricing system. The deviations reported in the subsequent tables represent the average differences across simulations within each tuple and counterfactual scenario.

On the demand side, I modify the indirect utility equation (Equation 9) by setting the cancellation fee, $f_{j,\Delta}$, equal to the price of the respective ticket, $p_{j,\Delta}$, thereby ensuring that no reimbursement is provided in the event of cancellation. On the supply side, adjustments depend on the counterfactual scenario.

First, in Equilibrium IV, which allows for reimbursement but prohibits re-selling, I modify the capacity transition equation (Equation 13) by setting the cancellation variable C_{fix} to zero, preventing capacity from being replenished. However, reimbursement costs ($f_j - p_j$) in the instantaneous revenue equation (Equation 12) are still accounted for. Second, in Equilibrium III, which permits re-selling but disallows reimbursement, the capacity transition equation remains unchanged, allowing capacity to be regained, while reimbursement costs ($f_j - p_j$) are set to zero.

7.1 No Cancellations

Table 5 presents the results of the counterfactual simulations based on the approach described above.

In the counterfactual scenario where cancellations are not permitted, discounted fixed ticket prices are, on average, 1.9% lower compared to the baseline. This indicates that the revenue manager internalises the absence of potential reimbursements across both ticket types, leading to delayed or forgone fixed ticket price increases. Combined with the discussion in Subsection 3.1, this evidence suggests that the marginal cost curve rotates only minimally when reselling is disallowed as prices are higher with cancellation than without.

Despite the lower (discounted) fixed ticket prices, consumer welfare declines in all four selected tuples when cancellations are removed. On average, consumer welfare decreases by 4.5% relative to the benchmark scenario with cancellations. Although the exogenous risk of cancellation is relatively low, the magnitude of this consumer welfare loss aligns with empirical findings on other dimensions of dynamic pricing in the existing literature. For instance, Williams (2022) reports an average consumer welfare decline of approximately 6.5% when an airline switches from uniform to dynamic pricing.

Firm revenue from both ticket types - fixed and flexible - also decreases in the absence of

Table 5: Counterfactual Results - No Cancellation

	(1)	(2)	(3)	(4)
Discounted Fixed Ticket Prices				
I - No Cancel, No Re-Sell	99.12%	97.20%	98.16%	97.97%
II - Yes Cancel, Yes Re-Sell	100.0%	100.0%	100.0%	100.0%
Consumer Welfare				
I - No Cancel, No Re-Sell	96.43%	95.46%	95.42%	94.72%
II - Yes Cancel, Yes Re-Sell	100.0%	100.0%	100.0%	100.0%
Revenue				
I - No Cancel, No Re-Sell	96.22%	93.07%	94.02%	93.86%
II - Yes Cancel, Yes Re-Sell	100.0%	100.0%	100.0%	100.0%
Route	A to B	B to C	A to C	B to A
Weekday	Friday	Sunday	Wednesday	Tuesday

Source: Author's own calculations using counterfactual equilibrium simulations.

Notes: Equilibrium I denotes a situation without cancellation and subsequent re-selling. Equilibrium II denotes the current situation, in which cancellation and subsequent re-selling is permitted. Changes in discounted fixed ticket prices are computed at the (simulation, day ahead)-level, while consumer welfare and revenue are summed across days ahead and then changes are computed at the simulation level. The individuals numbers reflect the average change in the respective statistic relative to Equilibrium II.

cancellations. Total ticket revenue declines by an average of 5.7%. Notably, the relative decline in revenue exceeds the relative drop in consumer welfare across all four tuples. Since discounted fixed ticket prices fall by less than the corresponding revenue decline, this suggests that some sales are lost⁹ This finding further supports the theory outlined in Subsection 3.1, as lower expected quantities are achieved without the option to cancel.

The counterfactual simulations conducted to decompose these net effects, corresponding to Equilibria III and IV, are presented in Table 6. The results indicate that the impact of cancellation is driven entirely by the demand side. Across the four selected tuples, the opportunity cost channel is found to have no meaningful effect, with changes in opportunity costs amounting to

⁹An alternative explanation is that consumers switch from the more expensive flexible ticket to the discounted fixed ticket.

Table 6: Counterfactual Results - No Cancellation Decomposition

	(1)	(2)	(3)	(4)
Discounted Fixed Ticket Prices				
III - No Cancel, Yes Re-Sell	99.12%	97.20%	98.16%	97.97%
IV - Yes Cancel, No Re-Sell	100.0%	100.0%	100.0%	100.0%
Consumer Welfare				
III - No Cancel, Yes Re-Sell	96.43%	95.46%	95.42%	94.72%
IV - Yes Cancel, No Re-Sell	100.0%	100.0%	100.0%	100.0%
Revenue				
III - No Cancel, Yes Re-Sell	96.22%	93.07%	94.02%	93.86%
IV - Yes Cancel, No Re-Sell	100.0%	100.0%	100.0%	100.0%
Route	A to B	B to C	A to C	B to A
Weekday	Friday	Sunday	Wednesday	Tuesday

Source: Author's own calculations using counterfactual equilibrium simulations.

Notes: Equilibrium III denotes a situation without cancellation but with re-selling. This equilibrium allows one to focus on the ticket quality channel. Equilibrium IV denotes a situation with cancellation but without re-selling. This equilibrium allows one to focus on the opportunity cost channel. Changes in discounted fixed ticket prices are computed at the (simulation, day ahead)-level, while consumer welfare and revenue are summed across days ahead and then changes are computed at the simulation level. The individuals numbers reflect the average change in the respective statistic relative to Equilibrium II.

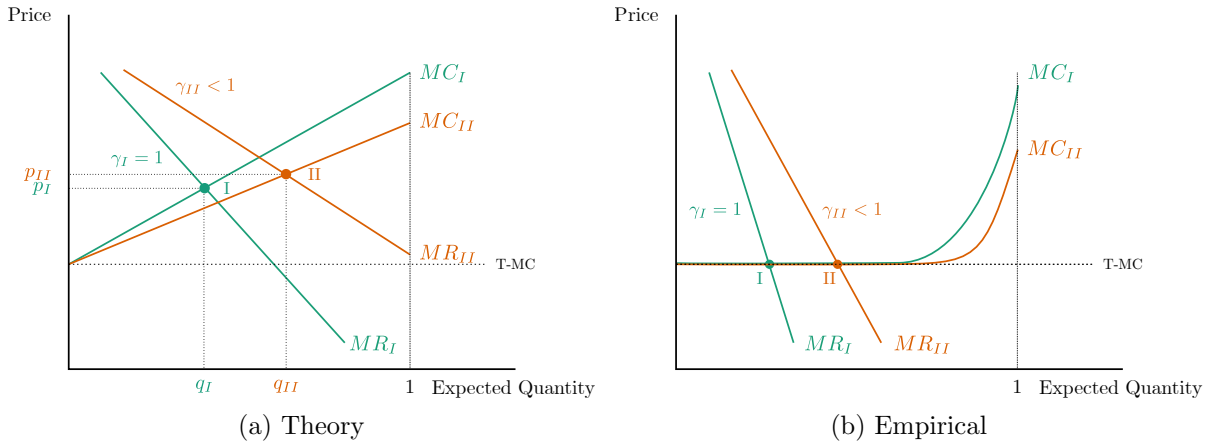
less than EUR 0.01 for the relevant quantity levels. To contextualise these findings, Figure 6a reproduces the theoretical framework, while Figure 6b provides a corresponding empirical illustration.

7.2 No Cancellations - Intertemporal

To evaluate the intertemporal effects of cancellations, I conduct the same counterfactual analyses as outlined in Subsection 7.1, with a focus on examining outcome variables across days prior to departure. Table 7 presents the intertemporal results comparing Equilibrium I outcomes to those of Equilibrium II.

Initial discounted fixed ticket prices, observed 22 to 28 days prior to departure, show minimal changes relative to their benchmark levels. This primarily reflects the fact that starting prices are

Figure 6: Theoretical Graph vs. Empirical Equivalent



Source: Authors' own illustrations.

Notes: Subscript *I* indicates the situation without cancellation, while Subscript *II* indicates a situation with cancellation. The marginal revenue curves are labelled MR, and the marginal cost curves are labelled MC. The horizontal line labelled T-MC denotes the technical marginal cost of providing the ticket. The variable γ indicates the share of the original price that the consumer must pay if she cancels her ticket, i.e. the cancellation fee. Please note that I have not explicitly drawn the discontinuity in MC_I , MC_{II} , $T - MC$ at $\mathbb{E}[q] = 0$. Panel (a) illustrates the theoretical prediction, and Panel (b) illustrates the empirical equivalent.

drawn from the empirical distribution and are not endogenously determined by the counterfactual revenue manager. Therefore, the resulting variation in fixed ticket prices arises mainly from the revenue manager forgoing price increases. As the day of departure approaches, discounted fixed ticket prices remain consistently lower despite increasing travel probabilities. The largest price declines, averaging 3.4%, are observed within the final seven days before departure.

Consumer welfare losses are most pronounced for early bookers. Compared to the baseline scenario with cancellations, consumer welfare for early bookers (22 to 28 days prior) decreases by an average of 23.2%. In contrast, late bookers (1 to 7 days prior) experience a more modest average welfare decline of 3.5%.

Intertemporal declines in firm revenue are less severe than the observed consumer welfare losses. Notably, the largest revenue drops occur 8 to 14 days before departure, with an average decrease of 8.5%. This period coincides with the arrival of the most price-sensitive consumers, particularly for the tuples represented in Columns (1) and (4).

8 Conclusion

The flexibility to cancel previously purchased items has become a prominent feature in many perishable good industries, including hospitality, passenger rail, and air travel. Using a theoretical framework, I demonstrate that offering the flexibility to cancel has an ambiguous effect on prices, and consequently on consumer welfare and firm revenue. This ambiguity arises from two competing channels: the higher quality channel and the opportunity cost channel. The higher quality channel reflects the increased utility for consumers, who can “insure” themselves against adverse events requiring cancellations, potentially leading to higher prices. Conversely, the opportunity cost channel arises from the firm’s ability to recover and re-sell previously sold capacity upon cancellation, thereby reducing the opportunity costs associated with selling today.

To empirically quantify the value of flexibility, I develop and estimate a structural dynamic pricing model and conduct counterfactual analyses. My findings indicate that disallowing cancellations leads to declines in both consumer welfare and firm revenue, by approximately 4.5% and 5.7%, respectively. An intertemporal analysis reveals that consumer welfare losses are particularly severe for early bookers, with declines of up to 35.6% compared to a scenario with cancellations, while late bookers experience an average welfare decline of only 3.5%. Overall, the results suggest that both consumers and firms benefit from allowing cancellations and the subsequent ability to re-sell capacity. In my particular setting, the opportunity cost channel has no meaningful impact¹⁰ and the observed effects are entirely attributable to the higher quality channel. This suggests that in contexts with higher opportunity costs, the losses from disallowing cancellations are likely to be more substantial.

¹⁰At relevant quantity levels, changes in opportunity costs are less than EUR 0.01.

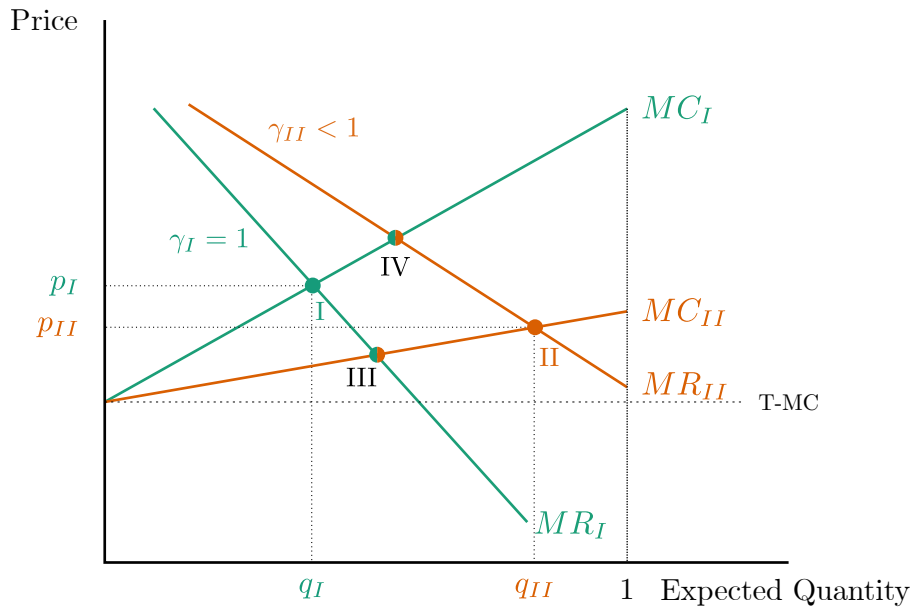
References

- Berry, Steven T. (1994). “Estimating Discrete-Choice Models of Product Differentiation”. In: *The RAND Journal of Economics* 25.2, pp. 242–262.
- Boston Consulting Group (2024). *How Companies Can Make Dynamic Pricing Fairer for Customers*. URL: <https://www.bcg.com/publications/2024/how-companies-can-make-dynamic-pricing-fairer-for-customers> (visited on 11/11/2024).
- Chen, Yanyou (Aug. 2024). “Network Structure and the Efficiency Gains from Mergers: Evidence from the U.S. Freight Railroads”.
- Cho, Sungjin, Gong Lee, John Rust, and Mengkai Yu (Apr. 2018). “Optimal Dynamic Hotel Pricing”.
- D’Haultfœuille, Xavier, Ao Wang, Philippe Février, and Lionel Wilner (Sept. 2023). “Estimating the Gains (and Losses) of Revenue Management”.
- Degiovanni, Pedro G. and Ron Yang (Nov. 2023). “Economies of Scale and Scope in Railroadings”.
- Hortaçsu, Ali, Olivia R. Natan, Hayden Parsley, Timothy Schwieg, and Kevin R. Williams (2023). “Organizational Structure and Pricing: Evidence from a Large U.S. Airline”. In: *The Quarterly Journal of Economics*.
- Lazarev, John (Jan. 2013). “The Welfare Effects of Intertemporal Price Discrimination: An Empirical Analysis of Airline Pricing in U.S. Monopoly Markets”.
- Porter, Robert H. (1984). “A Study of Cartel Stability: The Joint Executive Committee, 1880-1886”. In: *The Bell Journal of Economics* 14.2, pp. 301–314.
- Rust, John (1987). “Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher”. In: *Econometrica* 55.5, pp. 999–1033.
- Sampson, Hannah (2024). *Budget airlines are dropping notorious fees. What changed?* Accessed: 2024-11-11. URL: <https://www.washingtonpost.com/travel/2024/05/23/spirit-airlines-frontier-airlines-cheap-flights-cancellation-fees/>.
- Scott, Garrett (2022). “Pricing Dynamics and Screening Mechanisms”. PhD thesis. University of North Carolina at Chapel Hill.

- Sweeting, Andrew (2012). “Dynamic Pricing Behavior in Perishable Goods Markets: Evidence from Secondary Markets for Major League Baseball Tickets”. In: *Journal of Political Economy* 120.6, pp. 1133–1172.
- Vollmer, Drew (Feb. 2024). “Is Resale Needed in Markets with Refunds? Evidence from College Football Ticket Sales”.
- Williams, Kevin R. (2022). “The Welfare Effects of Dynamic Pricing: Evidence from Airline Markets”. In: *Econometrica* 90.2, pp. 831–858.

A Appendix

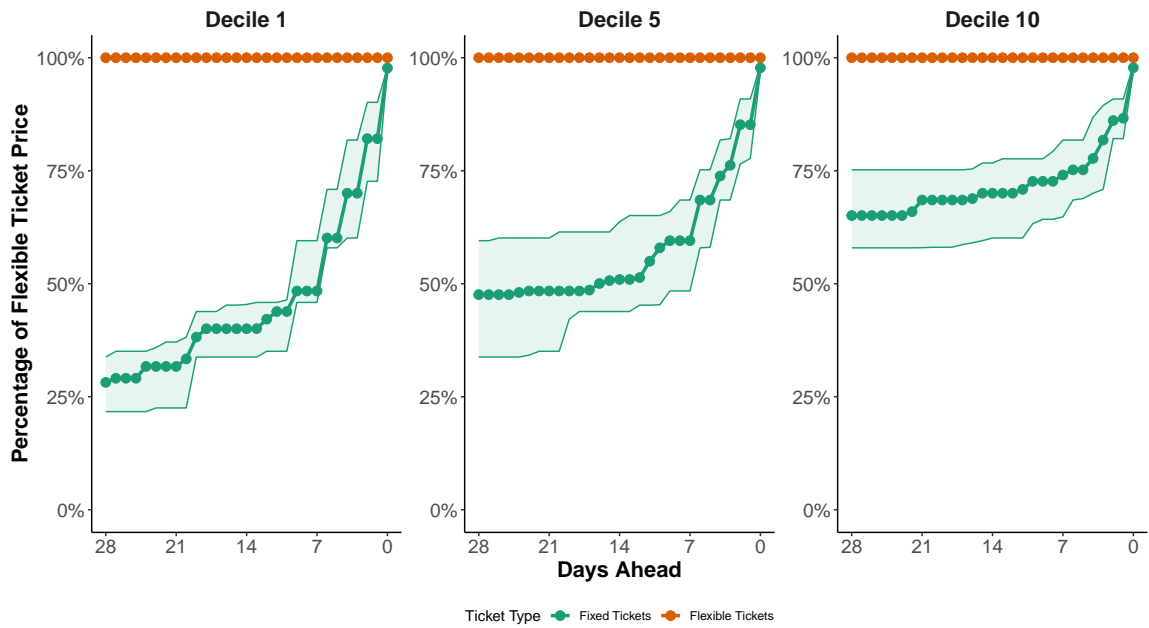
Figure A1: Supply and Demand in Situations Without and With Cancellation - Lower Price



Source: Authors' own illustrations.

Notes: Subscript I indicates the situation without cancellation, while Subscript II indicates a situation with cancellation. The marginal revenue curves are labelled MR, and the marginal cost curves are labelled MC. The horizontal line labelled T-MC denotes the technical marginal cost of providing the ticket. The variable γ indicates the share of the original price that the consumer must pay if she cancels her ticket, i.e. the cancellation fee. Please note that I have not explicitly drawn the discontinuity in MC_I , MC_{II} , $T - MC$ at $\mathbb{E}[q] = 0$.

Figure A2: Prices Across Days Ahead - By Decile of Expected Passengers



Source: Author's own calculations using TC-0 price and sales data on the subset of routes described in Subsection 2.2.

Notes: The x-axis represents the number of days before departure. The y-axis presents the price of the respective ticket type expressed as a percentage of the flexible ticket price. Conditional on a given number of days ahead, this percentage is computed for each (origin, destination, travel date, departure time)-tuple. Each dot along a solid line represents the median value of this measure's distribution. The lower and upper bounds of the shaded area represent the 25% and 75% percentiles of this measure's distribution, respectively. Each panel represents the observed pricing behaviour by decile of expected number of passengers on board the respective trains. The expected number of passengers is the predicted value from a linear regression of the number of scanned tickets onto a (train origin, stops, train destination, departure time)-specific fixed effect as well as a calendar day fixed effect.

Table 7: Counterfactual Results - No Cancellation Intertemporal

	(1)	(2)	(3)	(4)
Discounted Fixed Ticket Prices				
1-7 Days Ahead	98.31%	94.73%	97.40%	95.95%
8-14 Days Ahead	98.87%	96.32%	97.35%	97.56%
15-21 Days Ahead	99.78%	99.26%	98.86%	99.79%
22-28 Days Ahead	99.87%	99.94%	99.74%	99.97%
Consumer Welfare				
1-7 Days Ahead	97.50%	96.25%	96.15%	96.10%
8-14 Days Ahead	84.25%	92.76%	85.28%	89.09%
15-21 Days Ahead	82.17%	85.04%	79.52%	82.93%
22-28 Days Ahead	86.04%	64.41%	76.55%	80.06%
Revenue				
1-7 Days Ahead	98.01%	94.43%	95.18%	96.04%
8-14 Days Ahead	89.66%	93.69%	92.88%	89.77%
15-21 Days Ahead	93.53%	96.58%	97.75%	93.20%
22-28 Days Ahead	91.99%	93.39%	98.01%	91.96%
Route	A to B	B to C	A to C	B to A
Weekday	Friday	Sunday	Wednesday	Tuesday

Source: Author's own calculations using counterfactual equilibrium simulations.

Notes: Equilibrium I denotes a situation without cancellation and subsequent re-selling. Equilibrium II denotes the current situation, in which cancellation and subsequent re-selling is permitted. Changes in discounted fixed ticket prices are computed at the (simulation, days ahead group, day ahead)-level, while consumer welfare and revenue are summed across days ahead in the respective group and then changes are computed at the (simulation, days ahead group)-level. The individuals numbers reflect the average change in the respective statistic relative to Equilibrium II.

Table A1: Empirical Results - Daily Cancellation Risk of Fixed Tickets

	(1)	(2)	(3)	(4)
Risk of Cancellation				
$\tilde{\gamma}_1$	-4.417*** (0.091)	-3.958*** (0.084)	-4.123*** (0.106)	-4.147*** (0.113)
$\tilde{\gamma}_2$	-0.190 (0.026)	-0.215*** (0.026)	-0.166*** (0.034)	-0.211*** (0.032)
$\tilde{\gamma}_3$	0.005*** (0.001)	0.005** (0.001)	0.003 (0.002)	0.007*** (0.001)
Route	A to B	C to A	A to C	B to A
Weekday	Friday	Saturday	Wednesday	Tuesday

Source: Author's own calculations using TC-0 cancellation data for fixed tickets.

Notes: Standard errors in parentheses.

* ... $p < 0.05$ ** ... $p < 0.01$ *** ... $p < 0.001$

Table A2: Motivating Evidence - Cancellation as a Function of Purchasing Price

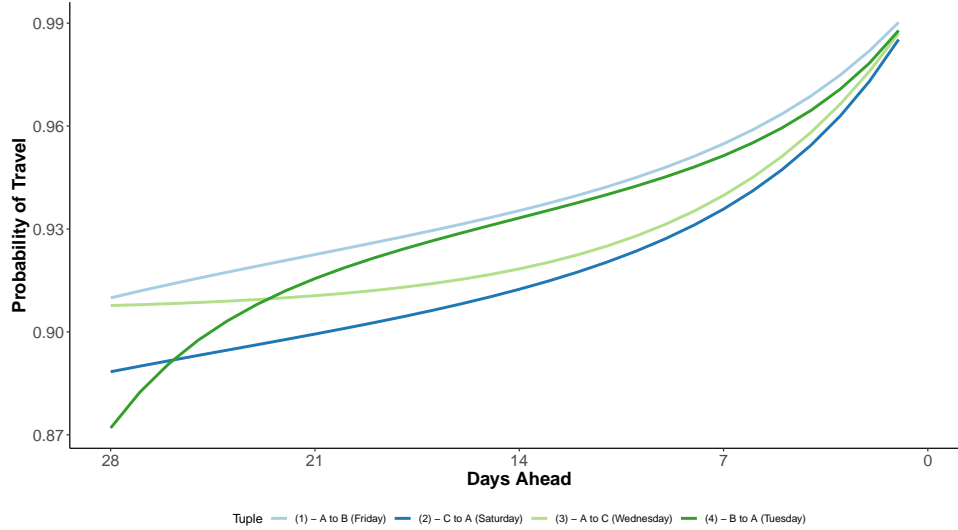
	$\mathbf{1}\{\text{Ticket Cancelled}\}$	
	(1)	(2)
Price	8.2E-04 (4.8E-04)	0.004 (0.003)
Fixed Effect	Origin \times Destination \times Weekday of Travel \times Days Ahead	Origin \times Destination \times Weekday of Travel \times Days Ahead
Method	Linear Probability Model	Logistic Regression

Source: Author's own calculations using TC-0 price, sales, and cancellation data for the subset described in Subsection 2.2. Only tickets, which were either cancelled or never validated, are used.

Notes: An observation is defined at the ticket-level. Standard errors are clustered at the (origin, destination, weekday of travel, days ahead)-level, and provided in parentheses. The dependent variable in both Columns 1 and 2 is equal to one if the respective ticket was cancelled before the cancellation deadline, and zero elsewhere. Column 1 features the linear probability model, estimated via OLS. Column 2 features the logistic regression model, estimated via maximum likelihood. In both columns, the coefficient of interest is on price. The fixed effect across both specifications is the interaction between the origin, destination, weekday of travel, and days ahead.

* ... $p < 0.05$ ** ... $p < 0.01$ *** ... $p < 0.001$

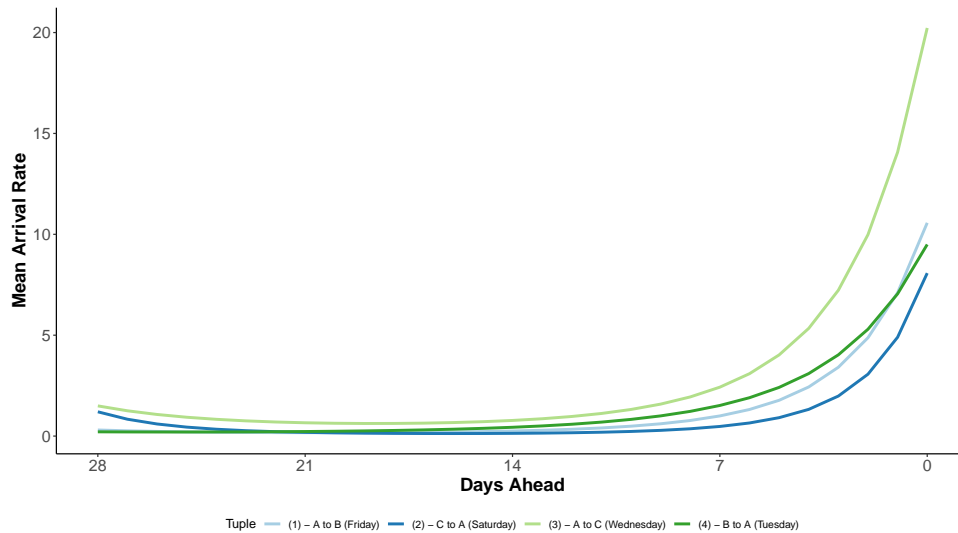
Figure A3: Probability of Travel by Days Ahead - μ_{Δ}



Source: Authors' own illustrations using the empirical cancellation risk estimates.

Notes: The x-axis indicates days prior to departure. The y-axis indicates the probability of travelling, conditional on booking on a certain day prior to departure. The four different coloured lines indicate the four selected tuples used to present the results.

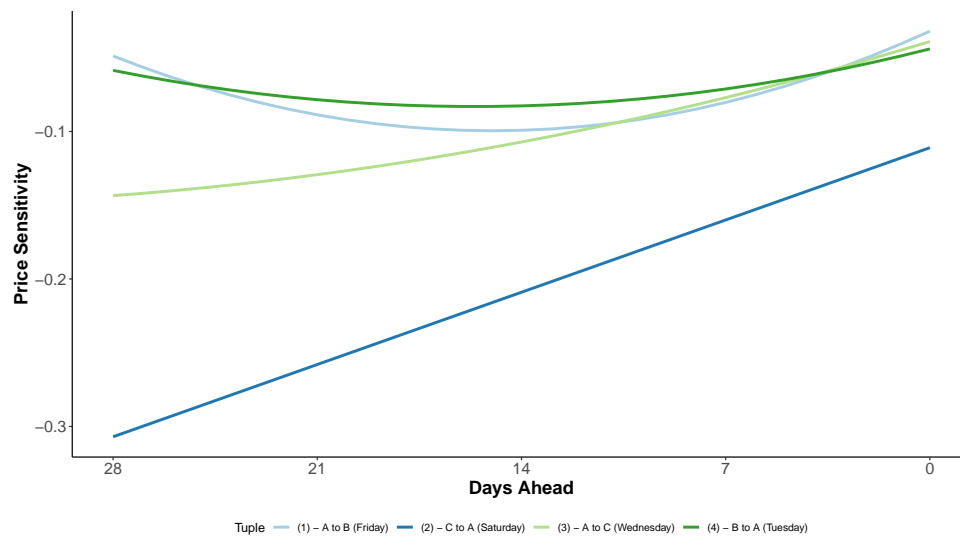
Figure A4: Mean of Potential Consumer Arrival Process by Days Ahead- λ_{Δ}



Source: Authors' own illustrations using the empirical arrival process estimates.

Notes: The x-axis indicates days prior to departure. The y-axis indicates the mean of the potential consumer arrival process, conditional on a certain day prior to departure. The four different coloured lines indicate the four selected tuples used to present the results.

Figure A5: Price Sensitivities by Days Ahead - α_{Δ}



Source: Authors' own illustrations using the empirical price sensitivity estimates.

Notes: The x-axis indicates days prior to departure. The y-axis indicates the price sensitivity parameter, conditional on a certain day prior to departure. The four different coloured lines indicate the four selected tuples used to present the results.