

Dynamic Oligopoly Pricing: Evidence from the Airline Industry*

Caspar Siegert

Robert Ulbricht

University of Munich

Toulouse School of Economics

November 18, 2015

Abstract

We explore how pricing dynamics in the European airline industry vary with competition and document patterns that are consistent with intertemporal price discrimination. First, the rate at which prices increase over time decreases in competition, supporting the idea that competition restrains the ability of airlines to price-discriminate against late-arriving customers. Second, the sensitivity to competition increases in the heterogeneity of the customer base, reflecting that restraints on price discrimination are only relevant when there is initial scope for discrimination. The patterns explain 83 percent of the observed within-flight price dispersion and 17 percent of the cross-market variation in pricing dynamics.

Keywords: Airline industry, dynamic oligopoly pricing, intertemporal price dispersion, price discrimination.

JEL Classification: D43, D92, L11, L93.

*We are extremely grateful to Anton Vasilev for invaluable help and technical support throughout the process of collecting our dataset. We also would like to thank Ken Boyer, Meghan Busse, Niko Matouschek, Mar Reguant, Klaus M. Schmidt, Mauricio Varela, and various seminar and conference participants for useful comments and discussions. Financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15 and GRK 801 is gratefully acknowledged. Finally, we like to thank Xueqian Chen and Markus Deák for research assistance. *Email Addresses:* caspar.siegert@bankofengland.co.uk, robert.ulbricht@tse-fr.eu.

1 Introduction

The tendency for airline ticket prices to rise as the scheduled departure date approaches is one of the most well-known regularities of dynamic oligopoly pricing. This tendency is often regarded as a prime example of intertemporal price discrimination, reflecting that customers in the airline industry are likely to have different demand elasticities correlating with their ability or willingness to book in advance.¹ While there is now a small literature with a focus on intertemporal price discrimination when firms have monopoly power (e.g., Lazarev, 2013; Williams, 2013), our understanding on the scope and working of intertemporal price discrimination in oligopoly markets is still limited. However, many important markets are oligopolistic and it is far from obvious how insights into intertemporal price discrimination from monopoly markets extend to oligopolies.

In this paper, we take a step towards filling this gap using new data on the time path of prices from the European airline industry. In particular, we empirically explore how pricing dynamics vary with the competitive environment and to what extent the identified variations are consistent with intertemporal price discrimination.

From an empirical perspective, the analysis of pricing dynamics has proved difficult mainly due to a lack of public data. In the airline industry, price data has been available at a (quarterly) level that does not allow to differentiate intertemporal price variations for a given flight (on a given date and time) from variations across different travel dates and from variations across different flights on a given route. This has led the empirical price discrimination literature to focus on the impact of competition on broader dispersion measures that do not differentiate among different dimensions of dispersion (e.g, Borenstein and Rose, 1994; Hayes and Ross, 1998; Gerardi and Shapiro, 2009). We address this issue using manually collected price data from the internet.² In particular, we construct a panel including about 1.4 million prices for airline tickets on the intra-European market where for each route-date pair (which we

¹The usual example are business travelers who learn about travel requirements at short notice and have a high willingness to pay for traveling at a particular time, and leisure travelers who have longer planning horizons and may be more flexible regarding their choice of time or even destination.

²Recent studies using similar approaches to collect airline price data include Lazarev (2013) and Williams (2013) who also explore pricing dynamics, but focus on monopoly markets. In contrast, McAfee and Te Velde (2007), Gaggero and Piga (2011), Escobari (2012), and Escobari, Rupp and Meskey (2013) have access to competitive price data, but are interested in the determinants of price levels and dispersion rather than pricing dynamics and their determinants.

refer to as a “market”) we record a time series of prices ranging from 10 weeks to 1 day prior to departure. Using this time series dimension permits us to shift the focus on pricing dynamics and their determinants.

We begin our analysis by estimating the intertemporal slope of prices and its sensitivity to competition. Overall, we find that prices in our sample increase substantially over time, but at a rate that is highly sensitive to competition. While monopoly prices increase by an average of 1.31 percent with every day that a customer waits to book, this slope is reduced to 1.19 percent in duopolies and continues to decrease monotonically to a slope of only 0.68 percent in markets with six competing airlines. A nonparametric treatment of pricing dynamics further reveals that these differences are mainly driven by the last 5 weeks before departure. Over this period, prices on monopoly routes increase by 95 percent, while prices on routes served by six competitors increase by only 50 percent.

At a first sight, these pricing dynamics are consistent with a theory of competitive price discrimination, reflecting the idea that competition restrains the ability of airlines to price-discriminate between customers with long planning horizons and customers who book at short notice. In an effort to make this argument more rigorous, we use hotel bookings data to construct measures for the tourist intensity in a given market and to proxy for variations in customer heterogeneity. If price discrimination is a major aspect of dynamic oligopoly pricing, then the degree to which an airline’s ability to price-discriminate becomes restrained as competition intensifies should depend on how much the airline would discriminate if it were unconstrained. Accordingly, if price discrimination plays an important role, competition should have a stronger impact on the intertemporal slope in markets with considerable customer heterogeneity.

Consistent with these considerations, we find that the sensitivity of the pricing dynamics to competition varies substantially with the identified heterogeneity. In markets characterized by a highly heterogeneous customer base, competition flattens the intertemporal slope from a daily rate of 1.42 percent in monopolies to 0.77 percent in markets with 6 competitors. In contrast, we find that in markets characterized by little customer heterogeneity, competition has virtually no impact on the intertemporal slope. The differential impact of customer heterogeneity is particularly noteworthy as we have to rely on noisy proxies for customer heterogeneity.

Comparing our results to the empirical price discrimination literature, the negative relation between competition and the intertemporal slope appears to be in line with the

finding of Gerardi and Shapiro (2009) that competition decreases dispersion (measured by the Gini coefficient on a given route across flights and travel dates), but it appears to be inconsistent with Borenstein and Rose (1994) who find the opposite. Using the ability to distinguish between the different dimensions of dispersion in our data, we show that the negative relation between competition and price dispersion is strongest when dispersion is defined as intertemporal within-flight dispersion; i.e., the dispersion in prices paid for the same physical flight. When dispersion is, in contrast, defined by pooling either across travel dates or across different flights that an airline operates on the same route, the relation becomes diluted (in the former case) or overturned (in the latter case). Hence, while we find competition to have an unambiguous negative impact on intertemporal price dispersion, the relation between competition and cross-flight dispersion in our data is less clear, which may explain seemingly contradicting findings in the earlier literature.

In the context of the empirical price discrimination literature, it is natural to ask: how much of the intertemporal price dispersion observed in our data is due to the pricing dynamics described above, and how much is due to unsystematic volatility within flights? We find that for the average flight the identified pricing dynamics account for 83 percent of the overall dispersion of prices for the same physical flight. That is, while there is a random component in prices that is consistent with stochastic demand fluctuations, a large share of the pricing dynamics that we observe follows a pre-determined path.

Looking at the *differences* in pricing dynamics across markets, about 17 percent of the observed variation can be explained by competition and our measures of customer heterogeneity. On the one hand, this underscores again the significant impact of competition and customer heterogeneity on pricing dynamics. On the other hand, this leaves a substantial amount of unexplained variation in pricing dynamics. This suggests that, while the impact of stochastic demand fluctuations on prices is small compared to the systematic trend of prices to increase, idiosyncratic fluctuations are likely to play a distinct role in explaining variations in trends across markets.³

The paper relates most closely to two areas of the literature: the empirical price dispersion literature and the literature on dynamic pricing schemes in oligopoly markets. In the context of the price dispersion literature (Borenstein and Rose, 1994;

³See also, Alderighi, Nicolini and Piga (2012) and Williams (2013) for empirical evidence that airlines dynamically adjust prices in response to demand fluctuations.

Hayes and Ross, 1998; Stavins, 2001; Gerardi and Shapiro, 2009), it is closely related to the works of Gaggero and Piga (2011) and Puller, Sengupta and Wiggins (2012), isolating specific dimensions of price dispersion to provide an explicit mapping to certain theoretical channels.⁴ Our paper provides a unifying perspective to these works by disentangling various dimensions of price dispersion and by linking them to an explicit identification of pricing dynamics.

Our investigation of pricing dynamics also complements a growing literature that examines dynamic pricing strategies in various industries.⁵ In the airline industry, Lazarev (2013) and Williams (2013) have recently used structural models to quantify the effects of price discrimination in monopoly markets, but do not consider changes in the market structure. At a methodological level, our reduced form identification approach is more closely related to Escobari (2012), who explores how prices adjust to demand shocks, and to Escobari, Rupp and Meskey (2013), who study price-discrimination between bookings in business hours versus bookings in the evening. To the best of our knowledge, this is the first paper which investigates empirically how (and why) pricing dynamics vary with the competitive environment.

The paper is structured as follows. Section 2 provides some theoretical background to fix ideas. Section 3 describes the data. Section 4 estimates the pricing dynamics and how they vary with the competitive environment. Section 5 introduces our customer heterogeneity measures and investigates the explanatory role of price discrimination. Section 6 explores how intertemporal price dispersion compares to various alternative dispersion categories. Section 7 explores the explanatory power of systematic pricing dynamics relative to idiosyncratic fluctuations, and Section 8 concludes.

⁴In particular, Puller, Sengupta and Wiggins (2012) study within-route price dispersion to empirically evaluate the “peak-load” pricing mechanisms formalized by Gale and Holmes (1992, 1993) and Dana (1998, 1999*a,b*, 2001). Similar to our analysis in Section 6, Gaggero and Piga (2011) look at several intertemporal price dispersion measures, but do not explore how these compare against other dimensions of dispersion and do not consider any of the dynamic pricing analysis which is at the core of our contribution.

⁵Outside the airline industry, Leslie (2004) and Courty and Pagliero (2012) have recently assessed intertemporal price discrimination in the market for Broadway theater and concert tickets. Similarly, Nair (2007) and Hendel and Nevo (2013) have recently explored intertemporal price discrimination for storable goods (video games and beverages, respectively).

2 Theoretical backdrop

Our exploration of the data is guided by the following hypotheses:

1. *Airline ticket prices increase as the scheduled departure date approaches.* We would expect to observe such a pattern if airlines price discriminate against late-arriving customers. Here the idea is that late-arriving customers are on average less price-elastic, reflecting, e.g., a high preference for traveling with a particular airline or at a particular time. Early customers on the other hand have longer planning horizons and may be more flexible regarding their choice of time or even destination.
2. *The increase in ticket prices should be less pronounced on markets that are subject to considerable competition.* The ability of airlines to exploit intertemporal differences in consumer preferences is restrained by competition. Accordingly, increases in prices over time should be largest for monopoly markets, and less pronounced in more competitive environments.
3. *The disciplining role of competition matters most in markets where the intertemporal customer heterogeneity is high.* When there are no meaningful differences between customers who book their flights early and customers who book their flights late, the impact of competition on the intertemporal slope in prices should be small as well.

In Appendix A, we provide a simple theoretical model that is based on Borenstein (1985) and generates these predictions. The model is based on the assumption that airlines offer differentiated goods (e.g., because flights differ with respect to the scheduled departure time, or because of certain brand preferences) and compete in prices on an early and a late market.

However, the basic intuition can also be conveyed by considering firms that offer an undifferentiated product and compete in quantities. Suppose there are N airlines, which compete in two distinct oligopoly markets: an “early” market E and a “late” market L . For simplicity, suppose airlines are fully symmetric and compete in quantities. In the standard Cournot equilibrium, the equilibrium markup in market $t \in \{E, L\}$ is given by

$$\frac{P_t - mc}{P_t} = \frac{1}{N\epsilon_t}, \quad (1)$$

where mc are marginal costs, P_t is the equilibrium price, and $\epsilon_t \equiv -\partial \log Q_t / \partial \log P_t$ is the market elasticity of demand. Precisely the same relation holds in the above-described differentiated goods environment that we explore in Appendix A.

The key feature both in the Cournot equilibrium and in the differentiated goods model in the appendix is that from the perspective of each airline, the *firm-specific* elasticity of demand, $N\epsilon_t$, increases in the number of competitors N , reflecting the usual logic that customers have more options to substitute across airlines when N is large. To fix ideas, suppose $N\epsilon$ is sufficiently large, so $\log(1 - 1/(N\epsilon)) \approx -1/(N\epsilon)$. Then from (1), the “intertemporal slope” can be expressed as:

$$\log\left(\frac{P_L}{P_E}\right) \approx \frac{1}{N} \left(\frac{1}{\epsilon_L} - \frac{1}{\epsilon_E} \right). \quad (2)$$

Using relation (2) we can easily derive the three testable predictions formulated above: (1) If late customers are less price-elastic than early customers, $\epsilon_L < \epsilon_E$, airlines charge higher prices in the late market than in the early market. (2) Due to the disciplining role of competition, the slope is decreasing in N .⁶ (3) The impact of N on the slope disappears when $\epsilon_E - \epsilon_L$ is small.

3 Data

We use a hand-collected panel dataset containing airline ticket prices on 92 intra-European routes and 41 distinct travel dates, where for each route-date pair we record a time series of prices ranging from 10 weeks to 1 day prior to departure. Figure 1 provides an overview over the cross-section of routes; details on the selection process and a full list can be found in Appendix B. Our dataset covers virtually all direct flights offered by traditional airlines on these routes, as well as most low cost carriers. We exclude indirect flights, as on intra-European routes these are arguably no close substitutes to direct flights.

Prices are recorded for flights taking off between October 31, 2010 and March 26, 2011, which defines the 2010/2011 European winter flight schedule.⁷ Within that period, we record prices for all flights leaving on Friday and returning on Sunday, as

⁶Here we implicitly treat elasticities as constant, but the conclusion is more general. Appendix A demonstrates this for the differentiated goods case where elasticities are proportional in prices.

⁷Flight schedules and routings within Europe are planned on a semiannual basis. Within these periods, routings and most flight characteristics do not typically vary.

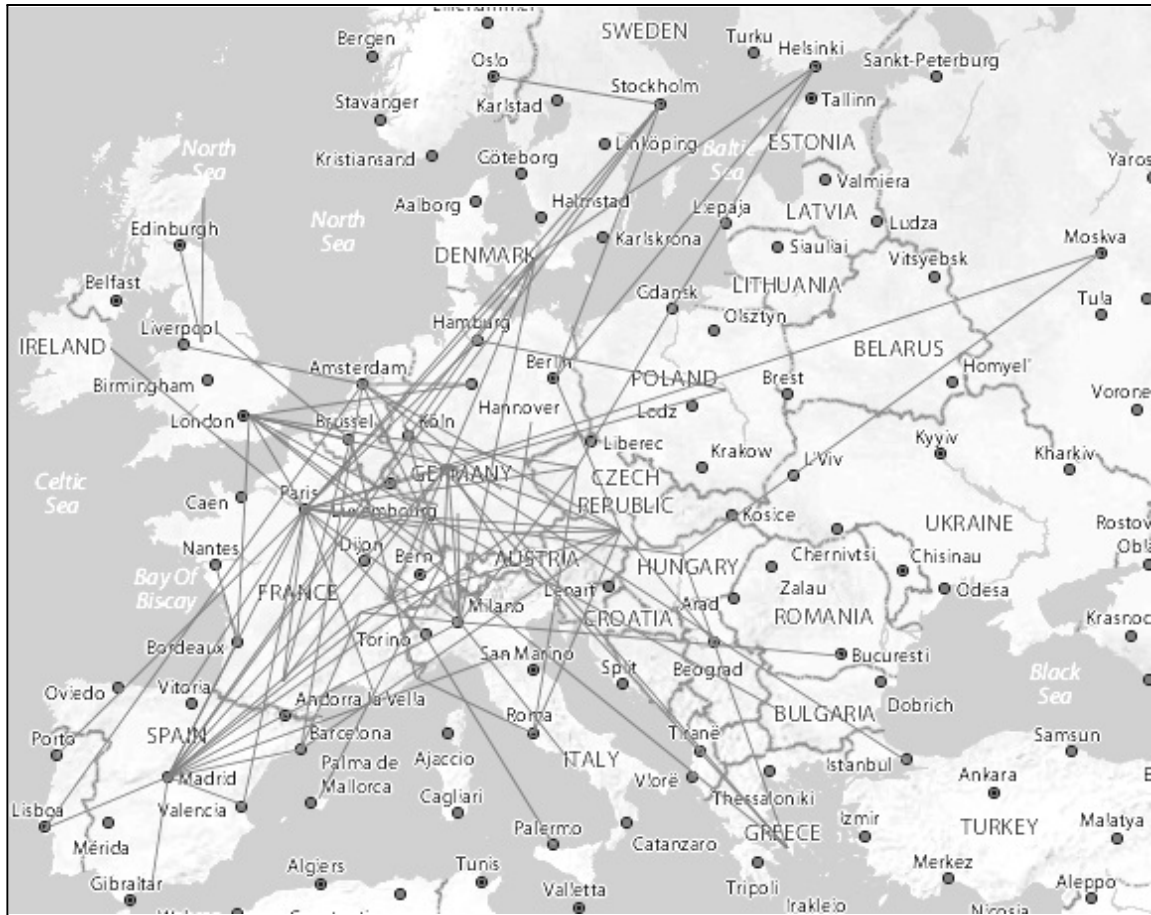


Figure 1: Map of routes

well as for all flights leaving on Monday and returning on Thursday. All recorded prices are for return tickets.⁸ This gives us two route-date pairs per route and week in our sample and 41 travel dates per route (uniquely defined by the date of the outbound flight). We refer to these route-date pairs as “markets”.

For each market, we record flights and prices once a week, starting 10 weeks prior to the travel date. In the last week prior to departure, prices are recorded on a daily basis to account for an increased frequency of price changes. Hence, we obtain up to 17 different prices for each physical flight.

In what follows, we will use the term “itinerary” to refer to a specific roundtrip itinerary, characterized by the combination of outbound and return flight identification

⁸We chose to focus on return tickets, since some European airlines still discriminate strongly against the purchase of one-way tickets. While the adoption (and recent abandoning) of such discrimination schemes poses interesting questions by itself, it is beyond the scope of this paper.

numbers.⁹ For example, in our terminology one “itinerary” on the route Paris–London would be using flight number BA 333 on the outbound flight and flight number BA 334 on the return flight. We reserve the term “flight” for the combination of an itinerary and a specific travel date. For each flight we record the cheapest available fare, treating code-sharing flights as distinct flights unless the involved airlines are affiliated through equity (see below).¹⁰

Overall we have data on 3762 out of 3772 distinct markets (41 travel dates times 92 routes).¹¹ Each market averages 377 prices that are recorded over up to 17 different dates prior to departure for an average of 41.9 flights (i.e., roundtrip combinations) per market. In total, our data set consists of 1.42 million individual prices (92 routes times 41 travel dates times 377 recorded prices per market). Routes are on average 560 miles long and connect metropolitan areas with an average of 3.9 million inhabitants. The share of domestic routes in our sample is roughly 13 percent (12 out of 92 routes).

Prices constitute offers by a leading website for airline ticket purchases, which accounts for a major share of bookings on the European market. The recorded prices in our sample range from 27 to 2581 Euros, with a weekly average of 364 Euros and a standard deviation of 466. Figure 2 relates the observed prices to the time remaining until departure. Prices increase from an average of 327 Euros ten weeks prior to departure to more than 500 Euros within the last week before departure.

To investigate the impact of competition on the observed pricing dynamics, we use the number of airlines that compete in a given market as our baseline measure of competition.¹² Thereby we treat airlines that are affiliated to each other as single competitors. An airline is matched to an affiliate group if that group owns more than 25% of the airline’s equity (see Appendix B for details). Table 1 summarizes the

⁹For each route, the direction of the outbound flight is randomly selected, so that for each city pair one of the two cities is the origin of the outbound flight, whereas the other one is the origin of the return flight.

¹⁰The reasoning behind this choice is that in so-called “block space” codeshare agreements, each codesharing partner is typically granted an *ex ante* fixed amount of seats with considerable freedom to set prices independently. In line with that, prices in our data differ substantially across different codesharers. See Appendix C.1 for details and further discussion, and for a robustness exercise where we only consider the cheapest available fare for a given physical flight and all its codesharing offers.

¹¹Of the ten remaining markets, seven are “missing” on the route Brussels–Leeds, where we did not find any flights offered on seven travel dates; the other three markets are missing on the routes Bordeaux–Madrid, Moscow–Budapest and Stockholm–Berlin.

¹²In 7.9 percent of our sample, the number of airlines offering services on the outbound leg differs from the number of airlines offering services on the return leg. In these cases, we set competition to the rounded up mean.

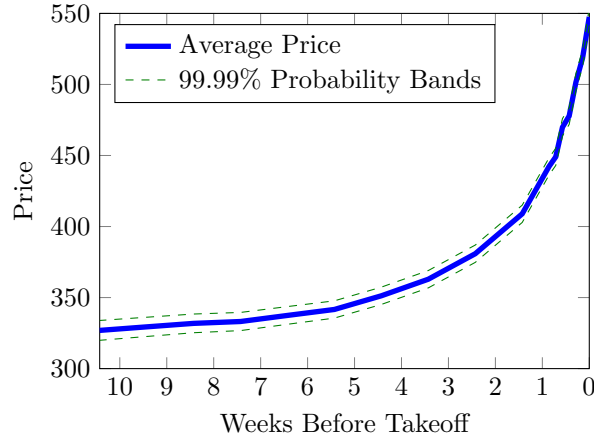


Figure 2: Average prices (in Euros) as a function of time remaining until departure

Table 1: Competition in the sample

Competing airlines	Prices		Markets	
	Frequency	Percent	Frequency	Percent
1	229 218	16.17	905	24.06
2	648 371	45.74	1696	45.08
3	275 680	19.45	656	17.44
4	185 051	13.05	382	10.15
5	68 237	4.81	107	2.84
6	11 078	0.78	16	0.43

resulting distribution of competitors. Alternative competition measures are explored in Appendix C.

4 Estimation of pricing dynamics

We begin our analysis by estimating the intertemporal slope of prices and its sensitivity to competition. Exploiting the within-flight time series structure of our data, our baseline empirical model estimates the intertemporal slope of log prices using only the intertemporal variations of prices *within* flights. All variations in prices that are route, time or itinerary-specific are absorbed by fixed effects specified for the travel date, the date of the price offer, and both the outbound and return itinerary (using flight numbers as identifiers). Section 4.2 generalizes our baseline setup to allow for nonlinear pricing dynamics and Appendix C provides some robustness analysis for

alternative competition measures.

4.1 Baseline specification

Let $Price_{ijtd}$ denote the price for a round trip that involves the outbound itinerary i and the return itinerary j (both identified by their flight numbers), for which the outbound flight departs at date t , and which is offered for sale at date d . Further, let $Comp_{ijt}$ denote a vector of dummy variables that covers all competition categories, and let $Daysleft_{td}$ denote the difference between t and d in days. As a baseline, we estimate the following equation:

$$\ln(Price_{ijtd}) = (\alpha + \beta Daysleft_{td}) \times Comp_{ijt} + \lambda_i + \mu_j + \nu_t + \xi_d + \varepsilon_{ijtd}, \quad (3)$$

where we treat λ_i , μ_j , ν_t , and ξ_d as fixed effects.¹³ Here, α is a vector of competition-specific constants and β is the relevant coefficient-vector on the interaction term $Comp_{ijt} \times Daysleft_{td}$. Note that λ_i and μ_j both nest a complete set of route specific fixed effects since any flight number uniquely pins down the corresponding city-pair. Together the specified set of fixed effects absorbs all itinerary-related effects such as departure time or length of flight; all route characteristics such as connected cities or alternative means of transportation; and all time-related effects such as travel dates, and dates of price offer.

The impact of competition on the observed pricing dynamics is captured by our estimates of β . Table 2 reports the estimated coefficients. Our estimates for the corresponding standard errors are adjusted for clustering at the market level. All reported coefficients are economically and statistically significant (at any reasonable level).¹⁴

It can be seen that for all competition categories prices increase as the scheduled travel date approaches, but that the intertemporal slope decreases substantially in the number of competitors. While in monopoly markets prices increase by an average of 1.31 percent with every day that a customer waits to book, this slope is reduced to

¹³Because our sampling is weekly for all but the last week before departure, a daily specification of fixed effects for d would completely absorb all last week effects of $Daysleft_{td}$; ξ_d is therefore modeled on a weekly level.

¹⁴The (unreported) competition-specific constants are only weakly identified in our sample by variations across travel dates but within routes since competition typically does not vary within routes for a given flight schedule. With this qualification in mind, we observe a hump-shaped relation between price levels and competition, peaking at three competitors.

Table 2: Baseline estimation of intertemporal slopes

	Coefficients	Clustered Std. Errors
$(Comp_{ijt} = 1) \times Daysleft_{td}$	-1.31	0.09
$(Comp_{ijt} = 2) \times Daysleft_{td}$	-1.19	0.09
$(Comp_{ijt} = 3) \times Daysleft_{td}$	-1.15	0.09
$(Comp_{ijt} = 4) \times Daysleft_{td}$	-1.07	0.09
$(Comp_{ijt} = 5) \times Daysleft_{td}$	-0.93	0.10
$(Comp_{ijt} = 6) \times Daysleft_{td}$	-0.68	0.12
Observations	1 417 635	
R-squared (adj.)	0.58	

Notes: The dependent variable is $\ln(Price_{ijt_d})$. Unreported but included in the estimation are levels of $Comp_{ijt}$ and fixed effects λ_i , μ_j , ν_t and ξ_d . Reported coefficients and standard errors are multiplied by 100. Standard errors are clustered at the market level. All reported coefficients are significant at the 0.1 percent level.

1.19 percent in duopoly markets and 0.68 percent in markets with 6 competitors.

These results are consistent with the idea that airlines operating in competitive markets are restrained in their ability to price-discriminate, leading to less intertemporal variation in prices. Section 5 explores this possibility further using hotel bookings data to identify variations in customer heterogeneity in an effort to make this argument more rigorous.

Note that throughout the paper, we treat competition as if it were exogenous. On the one hand, to the extent that competition on individual routes is driven in the short-run by the availability of landing slots or existing hubs, this assumption may be reasonable. On the other hand, endogenous variations in competition are unlikely to bias our estimates. We are interested in the slope of prices over time, and it is unclear how this slope would affect entry (considering that the route fixed effects and/or competition dummies control for any correlation between competition and average prices).

4.2 Nonlinear advanced purchase discounts

Our baseline estimation imposes the assumption that log prices depend linearly on $Daysleft_{td}$. Since this is likely not to be the case, we now generalize our baseline model to allow for nonlinear pricing dynamics. To this end, we use the same model as above, but replace $Daysleft_{td}$ by a vector of dummy variables covering *all* values of

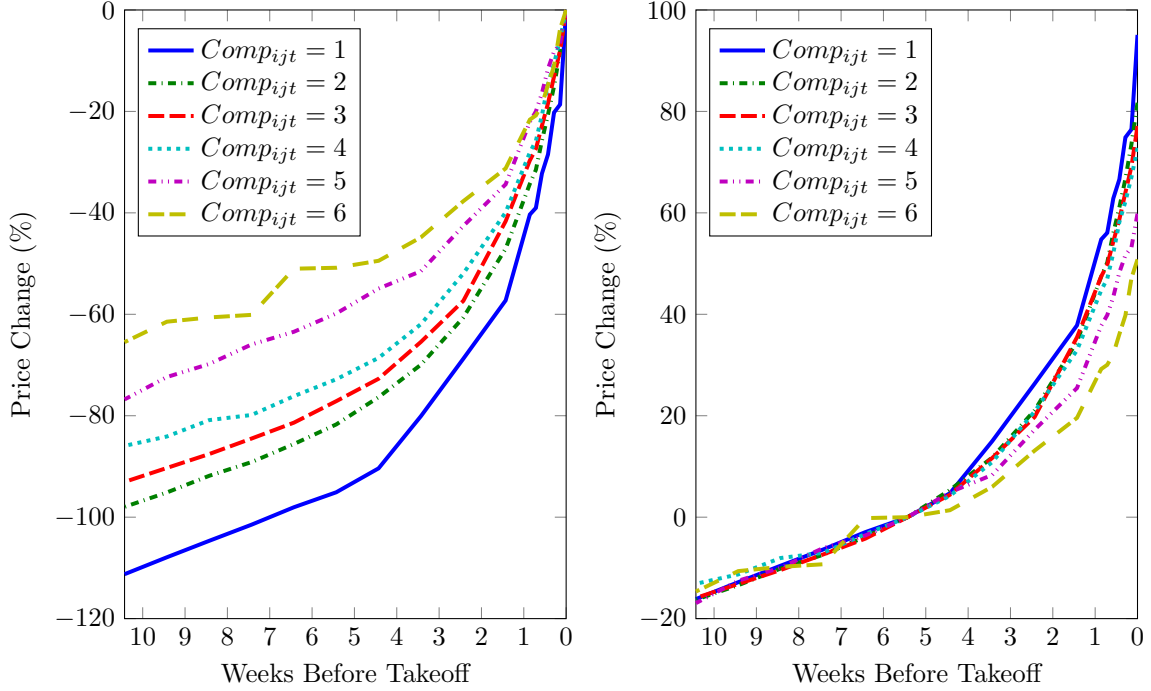


Figure 3: Nonlinear estimation of pricing dynamics.

Notes: Dependent variable is $\ln(\text{Price}_{ijt})$. Coefficients are multiplied by 100 and are normalized relative to the prices 1 day (left panel) and 38 days (right panel) before takeoff. Accordingly, the reported estimates approximately correspond to percentage price changes relative to the reported reference points. The estimation controls for levels of Comp_{ijt} and fixed effects λ_i, μ_j, ν_t and ξ_d .

Daysleft_{td} , hence imposing virtually no parametric restrictions on the evolution of prices over time. Our empirical model can be written as:

$$\ln(\text{Price}_{ijt}) = \phi(\text{Daysleft}_{td}, \text{Comp}_{ijt}) + \lambda_i + \mu_j + \nu_t + \xi_d + \varepsilon_{ijt}, \quad (4)$$

where $\phi : \{1, \dots, 74\} \times \{1, \dots, 6\} \rightarrow \mathbb{R}$ is an arbitrary function determined by the data.

The resulting nonlinear relationships are summarized in Figure 3 (see Figure 9 in the appendix for confidence bands). In the figure, we normalize ϕ such that for each Comp_{ijt} the average log price r days before takeoff is set to zero.¹⁵ The numbers on the y-axis report

$$\ln\left(\frac{p_{ijt}}{p_{ijt|d=t-r}}\right) \times 100,$$

¹⁵As before, we again scale all coefficients by 100. The displayed relations are hence given by $\{\phi(x, y) - \phi(r, y)\} \times 100$.

which approximately amounts to the difference in prices between date t and date r , expressed in percent of the price charged r days before departure.

In the left panel, we set $r = 1$. The y -axis thus approximately reflects the estimated advanced purchase discount relative to the price charged one day before departure. It can be seen that, although nonlinear, the slopes are again monotonically decreasing in the number of competitors. Accordingly, the relative discount for booking a flight in advance is less pronounced on routes that are served by a larger number of competitors, reinstating the conclusion drawn from our baseline estimation.

Taking a closer look at the identified pricing dynamics, it can further be seen that prices are increasing at similar slopes until about five weeks before takeoff. Only in the last five weeks, prices in less competitive routes have a significantly steeper slope than prices in more competitive routes. To illustrate this further, we set $r = 38$ in the right panel. Until about five weeks before takeoff log prices increase virtually along a single line across all competitive environments, showing an increase of about 0.45 percent for each day a customer waits to book. This translates into an overall discount of approximately 16% for purchasing tickets ten weeks before departure compared to five weeks before departure.

Starting about five weeks before departure, prices increase significantly faster, with prices on the least competitive environments increasing the fastest. On monopoly routes, customers pay a premium of 95.09 percent for purchasing their ticket a day before takeoff rather than five weeks in advance (the intercept with the y -axis on the right). This premium is reduced to 81.70 percent in duopoly markets, and is further reduced to 77.16 percent in markets with three competitors, 72.81 percent in markets with four competitors, 59.84 percent in markets with 5 competitors, and 50.83 percent in markets with 6 competitors.

In sum, the nonlinear estimation reinforces the clear pattern of dynamic oligopoly pricing revealed by our baseline specification: While airlines offer substantial advanced purchase discounts across all market structures, the magnitude of these discounts is highly sensitive to competition, suggesting that airlines operating in competitive markets are restrained in their ability to price-discriminate. The next section explores this possibility in more detail.

5 Intertemporal price discrimination

We now turn to the question whether price discrimination can explain the pricing dynamics identified in this paper. Above we have already argued that the observed “flattening” of the intertemporal slope for more competitive routes is in line with airlines being more restrained in their ability to price-discriminate. In this section, we construct a number of proxies for the heterogeneity in the customer base in a given market in an effort to make this argument more rigorous.

Following our argumentation in Section 2, the idea is that the degree to which an airline’s ability to price-discriminate becomes restrained in more competitive markets should depend on how much it would discriminate among customers if it were unconstrained (i.e., in a monopoly). In the extreme case where customers are completely homogeneous, we would not expect any price discrimination even in monopoly markets. Accordingly, if price discrimination is the driver behind the observed pricing dynamics, then competition should have less of an impact on the intertemporal slope in more homogenous markets.

5.1 Customer heterogeneity measures

In the airline industry the co-existence between tourists (and other leisure travelers) and business customers is arguably the largest source of heterogeneity in the customer base. Our approach to evaluating customer heterogeneity is therefore aimed at measuring the tourist intensity in a given market. In particular, we use Eurostat data on hotel bookings to construct two heterogeneity measures that are likely to be positively correlated with tourist intensity. In Appendix C we consider a few alternative measures, including dummies for itineraries that span Saturday-night stays, flights in the holiday season, and itineraries connecting cities with large temperature differentials. None of these alternative measures changes our qualitative results.

Our first measures ($Duration_{ij}$) computes the average number of nights per visit booked in hotels and similar residencies at the destination.¹⁶ Here the idea is that the average duration of a given booking is larger for tourists, so that small values of $Duration_{ij}$ are likely to reflect a high fraction of business travelers, whereas high

¹⁶The measure is computed using 2011 Eurostat data at the NUTS-2 level. For large cities the NUTS-2 level typically coincides with the city level (e.g., Berlin, Lisbon or Prague), while smaller cities or less densely populated areas are typically clustered into urban areas or regions (e.g., Manchester into Greater Manchester or Aberdeen into North Eastern Scotland).

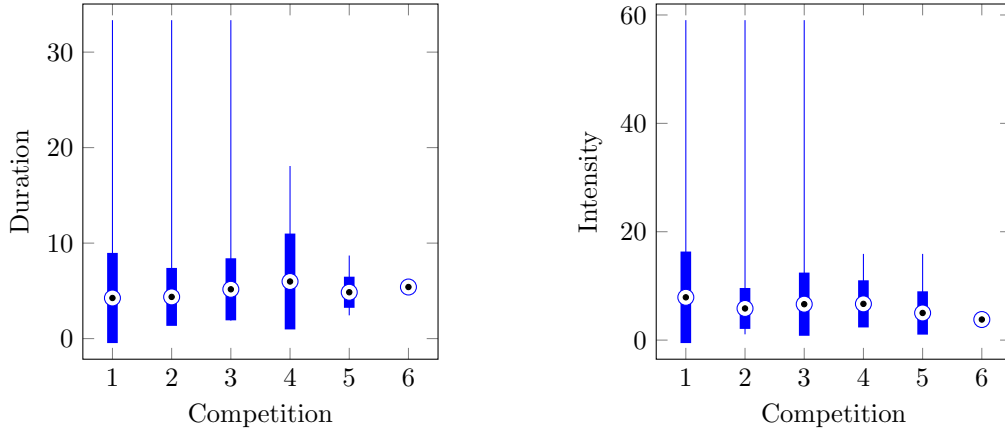


Figure 4: Summary of customer heterogeneity measures.

Notes: Means and ± 1 standard deviation bands are indicated by the circled dots and boxes. Supports are indicated by the extending lines.

values should reflect a larger share of tourists. Our second measure ($Intensity_{ij}$) is defined as the number of nights booked in hotels and similar residencies relative to the population at the destination (in thousands), capturing instead the intensity of the tourist industry at the destination.

To give an overview over these measures, Figure 4 plots their means (marked by the circled dots), standard deviation bands (marked by the boxes), and supports (marked by the lines) conditional on the number of competitors. It can be seen that the variation across competition categories is small relative to the within-category standard deviation and supports. This suggests that the competition effect that we have identified above are not merely driven by a correlation between customer heterogeneity and competition.¹⁷ Instead, even for a given (average) level of customer heterogeneity, airlines in more competitive markets seem to be less able to price discriminate.

Ex-ante, it is not obvious if the scope for price discrimination is increasing or decreasing in our specific measures of customer heterogeneity. This depends on whether on the average route, leisure travelers are an important source of revenues (in which case we would expect price discrimination to decrease as the share of leisure travelers

¹⁷More formally, the correlation coefficients are below 0.04 in absolute terms and are *not* statistical different from zero at the 1 percent level. Testing pairwise for difference in means between all 15 competition pairings for each measure, the data rejects any difference in 80 percent of the cases for $Duration_{ij}$ and in 67 percent of the cases for $Intensity_{ij}$.

increases further) or not (in which case we would expect the opposite). Given the scarcity of comprehensive statistics on the share of business versus leisure travelers, we let our data answer this question. As a robustness check we also include a specification that allows for non-monotonic effects: on routes where there is a small number of tourists, additional tourists may increase the heterogeneity of the customer base, while on markets with a large share of tourists, the same change may reduce heterogeneity (see Appendix C.3).

5.2 Impact of customer heterogeneity on pricing dynamics

To investigate how changes in customer heterogeneity affect the sensitivity of the intertemporal slope to competition, we estimate the following model:

$$\ln(\text{Price}_{ijt}) = (\alpha + \beta \text{Daysleft}_{td}) \times X_{ijt} + \lambda_i + \mu_j + \nu_t + \xi_d + \epsilon_{ijt}, \quad (5)$$

with

$$X_{ijt} = (1, N_{ijt}, \text{Het}_{ij}, N_{ijt} \times \text{Het}_{ij}),$$

where N_{ijt} is the number of competitors¹⁸, Het_{ij} is the customer heterogeneity measure, and λ_i , μ_j , ν_t and ξ_d are treated as fixed effects. To facilitate the discussion below, we normalize Het_{ij} to have zero mean and unit variance. Here β is the relevant vector of “slope” coefficients that define the intertemporal slope. Table 3 reports the estimated coefficients and standard errors (clustered at the market level).

From Panel (a) it can be seen that in line with our baseline estimation, the coefficients for Daysleft_{td} are negative (and large) in both specifications, whereas the coefficients for $\text{Daysleft}_{td} \times N_{ijt}$ are positive. Accordingly, in line with our baseline model, prices are again unambiguously increasing as the scheduled travel date approaches, and the intertemporal slope “flattens” in more competitive environments. Quantitatively, the identified competition effects are similar to our baseline model. Given that Het_{ij} has a zero mean and is largely uncorrelated with N_{ijt} this should not come as a surprise. Still, we find that customer heterogeneity has an important effect

¹⁸See Appendix C for a competition-specific treatment of customer heterogeneity. In sum, we do not find much evidence for non-linear effects of N_{ijt} . Accordingly, our conclusions are both qualitatively and quantitatively akin to the linear case. In an effort to streamline the following exposition, we hence focus on a linear model in this section and confine the competition-specific setting to the appendix.

Table 3: Competition, customer heterogeneity, and pricing dynamics

Customer Heterogeneity Measure:	(1) <i>Duration_{ij}</i>	(2) <i>Intensity_{ij}</i>
<i>(a) Estimated Slope Coefficients</i>		
<i>Daysleft_{td}</i>	−1.41 (0.10)	−1.37 (0.10)
<i>Daysleft_{td} × N_{ijt}</i>	0.07 (0.01)	0.08 (0.01)
<i>Daysleft_{td} × Het_{ij}</i>	−0.13 (0.02)	−0.16 (0.02)
<i>Daysleft_{td} × N_{ijt} × Het_{ij}</i>	0.06 (0.01)	0.08 (0.01)
<i>(b) Sensitivity and Overall Slope</i>		
Low heterogeneity markets		
sensitivity to competition	0.01	−0.01
implied monopoly slope	−1.27	−1.22
implied competitive slope (<i>N_{ijt} = 6</i>)	−1.22	−1.24
High heterogeneity markets		
sensitivity to competition	0.13	0.16
implied monopoly slope	−1.42	−1.38
implied competitive slope (<i>N_{ijt} = 6</i>)	−0.77	−0.60
Observations	1 255 125	1 286 687
R-squared (adj.)	0.57	0.57

Notes: The dependent variable is $\ln(\text{Price}_{ijt_d})$. Unreported but included in the estimations are levels of *Daysleft_{td}*, *N_{ijt}*, *Het_{ij}* and *N_{ijt} × Het_{ij}*, and fixed effects λ_i , μ_j , ν_t and ξ_d . The variable used for *Het_{ij}* is indicated at the top of each column. *Het_{ij}* is transformed to have zero mean and unit variance. Clustered standard errors (at the market level) are reported in parentheses. Reported coefficients and standard errors are multiplied by 100. All reported coefficients are significant at the 0.1 percent level. Low and high heterogeneity markets in Panel (b) are defined as markets with *Het_{ij}* = ±1 standard deviation relative to the mean. Number of observations depends on availability of heterogeneity data.

on pricing dynamics. The coefficients for *Daysleft_{td} × Het_{ij}*, are about twice the size of the ones for *Daysleft_{td} × N_{ijt}*, meaning that a one standard deviation change in customer heterogeneity has an impact on the intertemporal slope that is comparable to the effect of two additional competitors.

As discussed above, we need to take a stand on whether our tourist intensity measures increase or decrease the customer heterogeneity in a given market. We interpret a change in *Het_{ij}* that steepens the intertemporal slope in monopoly markets as an *increase* in customer heterogeneity. This is based on the presumption that

increases in the customer heterogeneity should steepen the intertemporal slope when airlines are unconstrained in their ability to price-discriminate. Given our empirical model, Het_{ij} steepens the slope in monopolies if

$$\frac{\partial^2 \ln(Price_{ijt_d})}{\partial Daysleft_{td} \partial Het_{ij}} \Big|_{N_{ijt}=1} = \beta_{Het} + \beta_{N \times Het} < 0,$$

where we use β_{Het} and $\beta_{N \times Het}$ to refer to the estimated slope coefficients for Het_{ij} and $N_{ijt} \times Het_{ij}$ in model (5). From the estimated coefficients, we can see that this is the case for both our measures, suggesting that Het_{ij} increases customer heterogeneity.

A potential concern regarding this interpretation of Het_{ij} is that changes in the tourist intensity may have a non-monotonic impact on customer heterogeneity. For instance, while heterogeneity is increasing in Het_{ij} on most routes, it may be decreasing in Het_{ij} on routes that are already characterized by a large number of leisure travelers. We explore this concern in Appendix C using a quadratic specification in Het_{ij} , but do not find any evidence for non-monotonicity. One potential explanation is that in our sample, we do not have a sufficient number of routes where leisure travelers are a dominant source of revenues to identify any decreasing effects of Het_{ij} on price discrimination. With this in mind, we henceforth interpret increases in Het_{ij} as increases in the customer heterogeneity.

Based on our interpretation of Het_{ij} , does the sensitivity to competition increase in customer heterogeneity as it should when price discrimination is the major source behind the observed pricing dynamics? Looking at the estimated slope coefficients of $N_{ijt} \times Het_{ij}$, it turns out that this is indeed the case: The estimated slope coefficient of $N_{ijt} \times Het_{ij}$ is statistically significant and positive in both specifications.

To assess the economic significance of this heterogeneity-caused variation in the sensitivity, we use the estimated model to compute the sensitivity for values of Het_{ij} at one standard deviation above and below its mean (meant to indicate markets with high and low customer heterogeneities, respectively). Exploiting the fact that we have normalized Het_{ij} to have zero mean and unit variance, these sensitivities are given by

$$\frac{\partial^2 \ln(Price)}{\partial Daysleft_{td} \partial N_{ijt}} \Big|_{Het_{ij}=\pm 1} = \beta_N \pm \beta_{N \times Het}.$$

Panel (b) of Table 3 reports the computed sensitivities and the implied overall slope. Driving down the heterogeneity to one standard deviation below its mean

virtually shuts down the sensitivity of the intertemporal slope to competition. This is consistent with the idea that in the absence of customer heterogeneity airlines do not discriminate much to begin with and, hence, they are also not getting restrained in their ability to price-discriminate when competition intensifies. Conversely, in high heterogeneity markets the slope of prices is highly sensitive to changes in competition. Here, an increase in competition “flattens” the intertemporal slope by between 0.13 and 0.16 percentage points per day.

To illustrate the magnitude of these effects, let us consider what this implies for the difference between a route that is served by a monopolist and a route that is served by six competitors. While in low heterogeneity markets the difference in the intertemporal slope between monopolies and competitive markets with $N_{ijt} = 6$ amounts to only 0.05 percentage points per day, the gap rises to 0.65 percentage points per day in high heterogeneity markets. We interpret the difference between these numbers as being likely to be caused by price discrimination. Compared to the overall slope in markets with $N_{ijt} = 6$, this gap amounts to nearly a doubling of the intertemporal slope, suggesting that intertemporal price discrimination is an important driver of the dynamic oligopoly pricing patterns that we have identified in this paper. The magnitude of the effects of customer heterogeneity are particularly noteworthy as customer heterogeneity is inherently hard to measure and we have to rely on a number of proxies that—at least individually—are likely to be rather noisy.

6 Decomposing price dispersion

To reconcile our study of pricing dynamics with the empirical price dispersion literature, we now construct a measure of intertemporal price *dispersion* and contrast it with more broadly defined dispersion measures used by the previous literature.

Due to limitations in the available data, the empirical price dispersion literature has so far focused on dispersion measures for airfares that do not disentangle intertemporal dispersion within flights from the dispersion in prices across different flights (pooling both, different itineraries and different travel dates). Different studies have thereby reached different conclusions regarding the relation between competition and these dispersion measures, ranging from positive (Borenstein and Rose, 1994), over no clear relation (Hayes and Ross, 1998), to negative (Gerardi and Shapiro, 2009).

We use the various dimensions of our panel to distinguish the impact of competition

on intertemporal price dispersion for a given flight from its impact on various types of cross-flight dispersion. Following the literature, we use the Gini coefficient to measure the dispersion within a given set of prices. Intuitively, the Gini coefficient corresponds to half the expected price difference in terms of the average price. A Gini coefficient of 0.10 would accordingly represent an expected absolute difference between two randomly selected prices of 20 percent of the average price.¹⁹

Panel (a) of Table 4 summarizes the average Gini coefficients in our data set. Column 1 reports the average price dispersion across all flights (and dates of ticket offers) offered by a given competitor on a given route. As in the previous literature, this pools together different itineraries (within the same route) and different flights across different travel dates. We find an average dispersion of 64 percent of the average offer.²⁰ Column 4 shows in contrast the intertemporal dispersion within a given flight (defined by a particular itinerary at a particular travel date). We see that the intertemporal dispersion is substantially smaller (24 percent).

The two intermediate cases where we pool flights only across itineraries (within markets) or only across travel dates (within itineraries) are reported in Columns 2 and 3. As expected, pooling along either dimension increases the dispersion relative to Column 4. The increase in dispersion due to differences in flight characteristics (e.g., departure time) appears to be higher than the dispersion due to different travel date characteristics (e.g., flights departing on Monday versus flights departing on Friday).

In the following we explore the relation between competition and the different dispersion measures. We do no longer include itinerary specific fixed effects, since these would absorb any dispersion on the itinerary and route level. Instead we include a large number of control variables X_{ijt} that were previously nested into λ_i and μ_j .²¹ Moreover, since the Gini coefficient is bounded between zero and one, we follow the previous literature and transform it into an unbounded statistic, using instead the Gini log-odds ratio $G_{ijt}^{lodd} = \ln[G_{ijt}/(1 - G_{ijt})]$.

¹⁹To make the Gini coefficients comparable to the previous literature, we compute all Gini coefficients after up-sampling observations with more than seven days left to departure in order to compensate for the increased sampling frequency in the last week.

²⁰Since the within-route dispersion is strongly right-skewed in our sample, the median dispersion is somewhat smaller, evaluating to 56 percent of the average offer.

²¹Specifically, X_{ijt} includes a full set of dummies for the departure city, destination city, and the time of the day of the outgoing and returning itineraries (at an hourly level); as well as measures for the market size, share, and symmetry; GDP and weather data of the connected cities; the weekday and length of the itineraries; holidays; and the tourist intensity in a given market (see Section 5 for a detailed presentation of these measures).

Table 4: Estimated and projected Gini coefficients

Dispersion Category:	Data			Projected Data		
	Routes (1)	Markets (2)	Itineraries (3)	Flights (4)	Flights (5)	Flights (6)
	<i>(a) Price dispersion within affiliate group \times dispersion category</i>					
	0.32 (0.13)	0.28 (0.14)	0.17 (0.09)	0.12 (0.09)	0.11 (0.04)	0.10 (0.03)
	<i>(b) Competition Effects on G_{ijt}^{load}</i>					
N_{ijt}	0.11 (0.01)	0.30 (0.03)	-0.22 (0.03)	-0.45 (0.07)	-0.54 (0.04)	-0.52 (0.04)
Observations	1 215 148	1 215 148	1 215 148	1 207 253	1 215 148	1 215 148
R-squared (adj.)	0.62	0.32	0.17	0.06	0.07	0.06

Notes: Panel (a) reports average Gini coefficients and standard deviations (in parenthesis) within routes (Column 1), markets (Column 2), itineraries (Column 3) and flights (Column 4). All categories are interacted with affiliate groups. Columns 5 and 6 of Panel (a) project within-flight Gini coefficients based on the pricing dynamics identified by the baseline regression (3) and the nonlinear regression (4). Panel (b) estimates the effect of competition on the Gini log-odds ratio, $G_{ijt}^{load} = \ln[G_{ijt}/(1 - G_{ijt})]$, in the data and in the projected sample. Unreported but included in the estimations are fixed effects ν_t and ξ_d and control variables listed in Footnote 21. Clustered standard errors (at the market level) are reported in parentheses. All reported coefficients are significant at the 0.1 percent level.

We estimate the following empirical model:

$$G_{ijt}^{lodd} = \alpha + \beta \times N_{ijt} + \gamma \times X_{ijt} + \nu_t + \xi_d + \epsilon_{ijt}, \quad (6)$$

where N_{ijt} is the number of competitors²², X_{ijt} contains our set of control variables, and ν_t and ξ_d are vectors of fixed effects for the travel date and the date of the price offer. The estimated coefficients are reported in Panel (b) of Table 4. All reported coefficients are significant at the 0.1 percent level (using standard errors that are clustered at the market level).

In line with the pricing dynamics identified above, competition has a negative impact on the intertemporal price dispersion (Column 4).²³ However, once we consider broader measures of price dispersion this effect is diluted (Column 3) or even overturned (Columns 1 and 2). A potential explanation is that in more competitive environments, airlines offer a larger number of connections to compete in departure dates and times, and that there is a meaningful dispersion of prices across these connections.²⁴

In sum, while we find an unambiguously negative impact of competition on intertemporal price dispersion, once we include cross-flight dispersion the impact depends on the precise definition of the dispersion measure. This suggest that the seemingly contradicting findings in the earlier literature may be driven by confounding different dimensions of dispersion.

7 Stochastic pricing dynamics

This section analyzes the relevance of the systematic pricing dynamics identified in Sections 4 and 5 relative to other sources of intertemporal price dispersion. While we do not attempt to explain these residual components, an obvious source consistent with the residual variations documented below are random demand fluctuations that cause airlines to dynamically adjust their prices.²⁵

²²Using the number of competitors N_{ijt} instead of a vector of dummy variables allows us to compare the impact of competition on the different dispersion measures in a straightforward way.

²³See also, Gaggero and Piga (2011).

²⁴The positive relation between competition and cross-flight dispersion is also consistent with theoretical explorations by Borenstein (1985) and Holmes (1989), who have shown that the dispersion *across* flights may increase in competition when consumers' cross-price elasticities between different airlines is lower than the elasticity of industry demand.

²⁵See, for instance, Alderighi, Nicolini and Piga (2012), Williams (2013), and Puller, Sengupta and Wiggins (2012). On the theory side, on particular channel why random demand fluctuations may

7.1 Systematic vs. unsystematic pricing dynamics

We start by asking how much of the intertemporal dispersion documented in the last section is due to the systematic advanced purchase discounts documented in Section 4 (instead of random price fluctuations). To answer this question, we use our estimates of the baseline model (3) and of the nonlinear model (4) to project samples of counterfactual price data in which the intertemporal variation in prices is fully determined by the identified intertemporal slopes. For these samples, we then compute the intertemporal dispersion on the flight level. We refer to this as the “systematic dispersion” implied by the estimated pricing dynamics. Columns 5 and 6 of Panel (a) in Table 4 report these intertemporal price dispersion measures in the two samples.

It can be seen that for the average flight the within-flight price dispersion projected from our baseline estimation (Columns 5 and 6) closely resembles the one in the data.²⁶ Based on the nonlinear estimation, 83 percent of the within-flight dispersion of the average flight (0.10 out of 0.12) can be attributed to the identified pricing dynamics seen in Figure 3. Systematic advanced purchase discounts therefore seem to be the major source of within-flight price dispersion, as opposed to unsystematic volatility around the trend.²⁷

We also use the projected price data to relate the (projected) Gini coefficient to the number of competitors. Using exactly the same specification as in the previous section, we find that the impact of competition on the dispersion that is due to systematic advanced purchase discounts is slightly higher (in absolute terms) than the impact on its empirical counterpart (Columns 5 and 6 in Panel (b) of Table 4 compared

cause variations in pricing dynamics in *ex ante* identical markets is modeled by the peak-load pricing literature (Gale and Holmes, 1992, 1993; Dana, 1998, 1999*a,b*, 2001).

²⁶The difference between the two estimations is likely due to the increased sampling frequency in the last week before departure. Because prices are generally increasing at a steeper slope during that period, the sample-average slope exceeds the time-average in our sample. Projecting the sample-average slope throughout the ten week horizon does therefore overestimate the contribution of systematic advanced purchase discounts relative to unsystematic volatility. This bias vanishes, once we allow slopes to vary with the time left until departure, as we do in the projection based on our nonlinear estimation.

²⁷Because the intertemporal slope maps nonlinearly to the projected Gini, our estimate is not identical to the one obtained from estimating the intertemporal slope for each flight and then using the flight-specific estimations to project the Gini. While such an exercise would be meaningless for the nonlinear estimation (for each flight, the identified nonlinear price path would be identical to the observed price path, so that the resulting Gini would perfectly predict the empirical one), it is feasible for the linear estimation. Doing so, we find a slightly higher average Gini of 0.12, suggesting that our results (slightly) underpredict the systematic contribution of pricing dynamics compared to random price fluctuations.

to Column 4). This suggests that the impact of competition on price dispersion is mainly driven by its impact on the intertemporal slope rather than changes in random fluctuations over time.²⁸

7.2 Systematic vs. unsystematic variations in pricing dynamics

Given that systematically increasing prices appear to account for a large share of the price dispersion on the average flight, we now address the related issue of how much of the *variation* in pricing dynamics across markets can be explained by the factors discussed above. In particular, we are interested in the combined explanatory power of competition and customer heterogeneity.²⁹

We proceed in two stages. In stage 1, we split our data into 3762 market-specific sub-samples, defined by the combination of a route r and a travel date t . In each of these market samples, we run the following first stage regression:

$$\ln(\text{Price}_{ijt}) = \alpha_{rt} + S_{rt}\text{Daysleft}_{td} + \epsilon_{ijt},$$

where α_{rt} and S_{rt} are the estimated market-specific coefficients. Collecting S_{rt} from all regressions, this gives us a sample of market-specific intertemporal slopes.

In stage 2, we then relate the identified slopes S_{rt} to a number of explanatory variables X_{rt} in order to assess their share in explaining the observed variations in S_{rt} . To this end, we estimate the following relation:

$$S_{rt} = \alpha + \beta \times X_{rt} + \epsilon_{rt}, \tag{7}$$

where X_{rt} varies across our specifications. For each specification, we use the R-squared statistic to identify the fraction of the cross-market variations in S_{rt} explained by X_{rt} . Table 5 reports the results.

In Column 1, we let $X_{rt} = \text{Comp}_{rt}$ to get an idea of the cross-market variations in the intertemporal slope that can be explained by variations in the competitive

²⁸In fact, consistent with the idea that the impact of competition on the empirical dispersion is primarily due to its impact on the systematic trend of prices to increase, the ratio of slopes $\frac{0.45}{0.52} \approx 0.865$ is very similar to the identified systematic dispersion share of 83 percent.

²⁹See Figure 8 in Appendix D for a schematic illustration how this differs from the exercise in the previous subsection.

Table 5: Explanatory power of customer heterogeneity

Part of X_{rt}	(1)	(2)	(3)
$Comp_{rt}$	Yes	Yes	Yes
$Comp_{rt} \otimes Het_{rt}$		Yes	Yes
$Comp_{rt} \otimes (Route_r, Flightdate_t)$			Yes
R-squared	0.05	0.17	0.63

Notes: The dependent variable is S_{rt} (the intertemporal slope in market rt , estimated by market-specific first stage regressions on $Daysleft_{td}$ and a constant). $Comp_{rt}$, $Route_r$ and $Flightdate_t$ are vectors containing a full set of competition, route, and flight date dummies, and Het_{rt} is a vector containing all herogeneity measures. The number of observations are 3311 (all markets with variation in $Daysleft_{td}$ and available heterogeneity data).

environment alone (analogous to the patterns identified in Section 4). We find that competition accounts for about 5 percent of the cross-market variation.

Given that our main interest lies on the share of cross-market variations explained by intertemporal price discrimination, we next let X_{rt} include a number of different market-characteristics that are likely to correlate with customer heterogeneity. Specifically, we include both measures used in Section 5 as well as a number of further measures that we discuss in detail in Appendix C. While individually each of these measures only captures some variations in customer heterogeneity, our hope is that their combination spans a large part of the actual variations in customer heterogeneity. We allow these heterogeneity measures to interact with competition in order to capture the effects discussed in Section 5. Using this specification, we can explain 17 percent of the variation in S_{rt} . This suggests that in addition to customer heterogeneity, there exist a number of unrelated factors that explain a large part of the variations in the intertemporal slopes across markets.

To get some idea at which level these factors are likely to operate, we lastly let X_{rt} include a set of route and travel date fixed effects (again interacted with $Comp_{rt}$). The explanatory power of this fixed effects specification is 63 percent, providing an upper bound on the explanatory power of *any* route or travel data specific characteristics. The remaining 37 percent are only explainable using factors that vary across travel dates *and* routes. While it is beyond the scope of this paper to identify the source of these residual variations in the intertemporal slope, we find it hard to reconcile them fully with any deterministic pricing scheme. We therefore conjecture that they represent, at least in parts, the adjustments of prices (and, hence, measured slopes) in

response to demand shocks that materialize within the last ten weeks before departure.

8 Summary

This study documents two sources of variations of airline pricing dynamics. First, the rate at which airline ticket prices increase as the scheduled departure date approaches is highly sensitive to the number of airlines active in a given market. While in monopoly markets prices increase by an average of 1.31 percent per day, this slope is reduced to 1.19 percent in duopoly markets, and is further decreasing in competition to 0.68 percent in markets with six competitors (the most competitive ones in our sample).

Second, this sensitivity of the intertemporal slope to competition varies substantially with the heterogeneity among customers. With a highly heterogeneous customer base, prices on monopoly routes increase by 0.65 percentage points per day more than prices on markets with six competitors. Conversely, on markets with very homogenous customers, this difference is almost zero.

We interpret these findings as evidence for intertemporal price discrimination being a major source of the observed pricing dynamics: Competition is likely to restrain an airline's ability to price-discriminate, but this constraint is more relevant in markets with an highly heterogeneous customer base where there is scope for price-discrimination in the first place.

Our results suggests that these forces are quantitatively important. Systematic advanced purchase discounts explain about 83 percent of the total within-flight dispersion on the average flight. Moreover, competition and customer heterogeneity explain 17 percent of the observed variation in the intertemporal slope across markets.

These numbers suggest that while systematically increasing prices are the main source behind intertemporal price dispersion, there are likely a number of factors alongside price discrimination that are responsible for the observed variations in the intertemporal slope. In particular, we find that 37 percent of the cross-market variation can only be accounted for by factors that vary across travel dates *and* routes. While it is beyond the scope of this paper to identify the source of these residual variations, we conjecture that price adjustments in response to random demand fluctuations are likely to play an important role in explaining these residual variations.

References

- Alderighi, Marco, Marcella Nicolini, and Claudio A Piga.** 2012. “Combined effects of load factors and booking time on fares: Insights from the yield management of a low-cost airline.”
- Borenstein, Severin.** 1985. “Price discrimination in free-entry markets.” *Rand Journal of Economics*, 16(3): 380–397.
- Borenstein, Severin, and Nancy L Rose.** 1994. “Competition and Price Dispersion in the US Airline Industry.” *Journal of Political Economy*, 102(4): 653–683.
- Brueckner, Jan K, Nichola J Dyer, and Pablo T Spiller.** 1992. “Fare determination in airline hub-and-spoke networks.” *The Rand Journal of Economics*, 309–333.
- Courty, Pascal, and Mario Pagliero.** 2012. “The Impact of Price discrimination on Revenue: Evidence from the concert industry.” *Review of Economics and Statistics*, 94(1): 359–369.
- Dana, James D Jr.** 1998. “Advance-purchase discounts and price discrimination in competitive markets.” *Journal of Political Economy*, 395–422.
- Dana, James D Jr.** 1999a. “Equilibrium price dispersion under demand uncertainty: the roles of costly capacity and market structure.” *The RAND Journal of Economics*, 30(4): 632–660.
- Dana, James D Jr.** 1999b. “Using yield management to shift demand when the peak time is unknown.” *The RAND Journal of Economics*, 30(3): 456–474.
- Dana, James D Jr.** 2001. “Monopoly price dispersion under demand uncertainty.” *International Economic Review*, 42(3): 649–670.
- Escobari, Diego.** 2012. “Dynamic Pricing, Advance Sales and Aggregate Demand Learning in Airlines.” *The Journal of Industrial Economics*, 60(4): 697–724.
- Escobari, Diego, Nicholas Rupp, and Joseph Meskey.** 2013. “Dynamic price discrimination in airlines.” *mimeo*.

- Gaggero, Alberto A, and Claudio A Piga.** 2011. “Airline market power and intertemporal price dispersion.” *The Journal of Industrial Economics*, 59(4): 552–577.
- Gale, Ian L, and Thomas J Holmes.** 1992. “The efficiency of advance-purchase discounts in the presence of aggregate demand uncertainty.” *International Journal of Industrial Organization*, 10(3): 413–437.
- Gale, Ian L, and Thomas J Holmes.** 1993. “Advance-purchase discounts and monopoly allocation of capacity.” *The American Economic Review*, 135–146.
- Gerardi, Kristopher S, and Adam Hale Shapiro.** 2009. “Does competition reduce price dispersion? New evidence from the airline industry.” *Journal of Political Economy*, 117(1): 1–37.
- Goolsbee, Austan, and Chad Syverson.** 2008. “How do incumbents respond to the threat of entry? Evidence from the major airlines.” *The Quarterly Journal of Economics*, 123(4): 1611–1633.
- Hayes, Kathy J, and Leola B Ross.** 1998. “Is airline price dispersion the result of careful planning or competitive forces?” *Review of Industrial Organization*, 13(5): 523–541.
- Hendel, Igal, and Aviv Nevo.** 2013. “Intertemporal Price Discrimination in Storable Goods Markets.” *The American Economic Review*, 103(7): 2722–2751.
- Holmes, Thomas J.** 1989. “The effects of third-degree price discrimination in oligopoly.” *American Economic Review*, 79(1): 244–250.
- Lazarev, John.** 2013. “The Welfare Effects of Intertemporal Price Discrimination: An Empirical Analysis of Airline Pricing in US Monopoly Markets.” *mimeo*.
- Leslie, Phillip.** 2004. “Price discrimination in Broadway theater.” *RAND Journal of Economics*, 520–541.
- McAfee, R Preston, and Vera Te Velde.** 2007. “Dynamic pricing in the airline industry.” *Handbook on Economics and Information Systems*.

- Nair, Harikesh.** 2007. “Intertemporal price discrimination with forward-looking consumers: Application to the US market for console video-games.” *Quantitative Marketing and Economics*, 5(3): 239–292.
- Puller, Steven L, Anirban Sengupta, and Steven N Wiggins.** 2012. “Does scarcity drive intra-route price dispersion in airlines?” *NBER Working Paper*, 15555.
- Stavins, Joanna.** 2001. “Price discrimination in the airline market: The effect of market concentration.” *Review of Economics and Statistics*, 83(1): 200–202.
- Williams, Kevin R.** 2013. “Dynamic Airline Pricing and Seat Availability.” *mimeo*.

A Intertemporal price discrimination in a differentiated goods environment

In this appendix, we demonstrate how the relation between markups, the demand elasticity, and competition in equation (1) can also be derived in a simple differentiated goods environment. The model is based on Borenstein (1985). Assume there are two periods, an “early” period E and a “late” period L : $t \in \{E, L\}$. In each period, there is a number of different standard Salop markets indexed by i . Each of these markets can be represented as a circle with circumference of one.

Any firm that offers flights is assumed to be active on all markets; i.e., the number of competitors $N \geq 2$ is constant across periods and markets. Firms are indexed by j and are located equidistantly along the circle. Firms incur a marginal cost of m when selling a ticket. In order to focus on intertemporal price discrimination, we rule out the possibility for airlines to discriminate intra-temporally across different markets. Formally, airlines can observe the period in which a customer buys his ticket, but cannot distinguish between customers located on different markets i . Hence, each airline j is restricted to offering two prices, P_{jE} and P_{jL} .

In each period, there is a total mass one of customers, a fraction θ_{it} of which are located on market i . On each of these markets, consumers are uniformly distributed along the circle, with $z \in [0, 1]$ denoting a consumer’s location on the circle. Consumers have an outside option of zero and receive a utility of $U_{it} = A_i - P_{jt} - c_i|z - X|$ when buying a ticket from firm j , where X is the location of firm j and $|z - X|$ is the arc distance from z to X . Different markets differ with respect to consumer’s willingness to pay for a ticket A_i and the intensity of their “product preferences” c_i . In the context of the airline industry, one can think of c_i , for instance, as the value of flying at a particular time with a particular airline. In that sense, we can think of different markets as consisting of different types of consumers (e.g., business versus leisure travelers), and the distribution of consumer across types may differ in the two periods.

We assume that $A_i \in [\underline{A}, \bar{A}]$ and $c_i \in [\underline{c}, \bar{c}]$ and denote the distribution of c_i over consumers who buy in period t by $F_t(c)$. In order to keep the analysis tractable and to focus on the most interesting case, we impose two parameter restrictions. First, we assume that \underline{c} is sufficiently large such that a symmetric equilibrium exists.³⁰

³⁰A (grossly) sufficient condition for this to be the case is that $\underline{c} \geq 2/3 [E_t(1/c_i)]^{-1}$ for all $t \in \{E, L\}$.

Moreover, we assume that \underline{A} is sufficiently large relative to m such that in equilibrium, all consumers purchase a ticket and firms compete with each other for customers on all markets.³¹ Since, the airline industry is typically characterized by high fixed costs and low marginal costs, it seems likely that in practice the willingness to pay A_i is indeed high relative to the marginal cost m and it is optimal to serve all customers who are willing to buy a ticket.

The optimization problems in the two periods are independent of each other, so we can solve them separately from each other. Assume there is a symmetric equilibrium in which all consumers find it strictly optimal to purchase a ticket and airlines charge price P_t . Consider the number of tickets that firm $j = 1$ (which is located at an arbitrary point 0 on the circle) sells on market i in period t . The marginal customer with $z \geq 0$ who buys from firm 1 must be indifferent between buying from airline 1 or airline 2: $A_i - P_{1t} - c_i z = A_i - P_t - c_i(1/N - z)$, or

$$z = \frac{1}{2} \left(\frac{P_t - P_{1t}}{c_i} + \frac{1}{N} \right).$$

Observe that firm 1 does not only serve customers located within the interval $z \in [0, 1/N]$ but also on its other side; i.e., on $[-1/N, 0]$. Summing over all markets, the total amount of customers that are served by firm 1 is therefore given by

$$Q_{1t}(P_{1t}, P_t) = \sum_i \left(\frac{P_t - P_{1t}}{c_i} + \frac{1}{N} \right) \frac{1}{\theta_{it}}.$$

In equilibrium, firm 1 chooses P_{1t} in order to maximize profits

$$\Pi_{1t} = (P_{1t} - m)Q_{1t}(P_{1t}, P_t).$$

Imposing symmetry, the first-order condition can be written as

$$\frac{P_t - m}{P_t} = \frac{1}{N\epsilon_t},$$

where $\epsilon_t = \mathbb{E}_t\{1/c_i\}P_t$ and $N\epsilon_t = -\partial \log Q_{1t} / \log P_{1t}$ is firm 1's elasticity of demand. Given the constraints on A_i and c_i , this is indeed an equilibrium and the symmetric

³¹A sufficient condition for all markets to be covered is that $\underline{A} > m + 1.5\bar{c}/N$. Without this assumption, the problem becomes highly intractable (see Borenstein, 1985 for details).

equilibrium is unique.³² That is, just as in the standard Cournot case, the firm’s elasticity of demand scales with N and markups are given by relation (1).

In the context of the differentiated goods setting, airlines charge higher prices if customers have stronger (and less uncertain) “product preferences”. If on average, late customers have more intensive product preferences (e.g., because they want to travel with a specific airline, on a specific day or at a specific time), prices increase over time.³³ However, the extent to which firms can exploit consumers’ product preferences and set prices in excess of their marginal costs is decreasing in competition. Hence, competition restrains the ability of airlines to price-discriminate between customers who book their flights early and customers who book at short notice. As a result, the increase in prices over time is less pronounced in more competitive environments. To see this formally, rewrite the first-order condition to get $P_t = m + 1/(N\tilde{\epsilon}_t)$ where $\tilde{\epsilon}_t = \epsilon_t/P_t = \mathbb{E}_t\{1/c_i\}$ captures the exogenous variation in price elasticities. Taking logs, we get

$$\log\left(\frac{P_L}{P_E}\right) = \log\left(\frac{\tilde{\epsilon}_L^{-1} + Nm}{\tilde{\epsilon}_E^{-1} + Nm}\right),$$

which is clearly decreasing in N .

Finally, these effects become less relevant if there are no meaningful differences between customers who book their flights early and customers who book their flights late (i.e., if $\epsilon_E - \epsilon_L$ is small). In this case firms would not discriminate between different types of customers in any meaningful way even if they were unconstrained. Hence, competition has a very limited effect.

It is worth noting that our results do not hinge on the independence of demands across the two markets. In particular, it should be clear that given that prices increase over time, customers that arrive early would have no incentives to postpone buying a ticket if we would allow for such actions.

³²More specifically, we can show that in the candidate equilibrium, all consumers do indeed find it optimal to purchase a ticket, and that there are no non-local deviations that firms may find profitable. Finally, there is no symmetric equilibrium in which some consumers do not buy a ticket.

³³Specifically, we have that $\epsilon_L < \epsilon_E$ whenever $F_L(c)$ has second-order stochastic dominance over $F_E(c)$.

B Data construction

B.1 Routes

Our cross-section of routes is sampled from the existing connections between a set of 60 European cities with international airports. All routes are defined on the city-level. In case there exist multiple airports within one city, we include routes to all airport combinations (e.g., routes between London and Paris cover all offered combinations between {LCY, LGW, LHR, LTN} and {CDG, ORY}).

The 60 cities were chosen to ensure regional variety as well as variety in the size and importance of the residing airports. To this end, the set includes the cities with the four largest airports in each of the EU5 countries (measured by 2009 total passenger traffic):

- France: Paris, Nice, Lyon, Marseille
- Germany: Frankfurt, Munich, Duesseldorf, Berlin
- Italy: Rom, Milan, Venice, Catania
- Spain: Madrid, Barcelona, Palma de Mallorca, Malaga
- UK: London, Manchester, Edinburgh, Birmingham

The remaining 40 cities are selected from both the EU5 and the rest of Europe (including Russia and Turkey): Aberdeen, Amsterdam, Athens, Belgrade, Bilbao, Bologna, Bordeaux, Brussels, Bucharest, Budapest, Copenhagen, Dublin, Geneva, Hamburg, Hannover, Helsinki, Innsbruck, Istanbul, Leeds, Leipzig, Lisbon, Liverpool, Moscow, Nantes, Naples, Nuernberg, Oporto, Oslo, Palermo, Prague, Sofia, Stockholm, Strasbourg, Stuttgart, Toulouse, Turin, Valencia, Vienna, Warsaw, and Zurich.

From the simplex of routes spanned by those 60 cities, we then sampled 100 random routes, disregarding all routes for which not at least one direct daily connection was offered at the beginning of our sampling period (October 31, 2010). To this routes, we added, if not yet contained, the ten routes connecting the cities with the largest airports in each of the EU5 countries (Paris, Frankfurt, Milan, Madrid, and London).

A limitation of our data source is that it does not contain prices set by Ryanair, a major competitor in the intra-European market. To prevent our data from being affected by an unobserved competitor, we therefore excluded all routes that were

served by Ryanair within a 40 miles (65 km) radius of the corresponding city centers. From the above city list, this applies to all (existing) Ryanair route combinations between the following cities: Barcelona, Birmingham, Berlin, Bologna, Bordeaux, Brussels, Budapest, Catania, Dublin, Edinburgh, Leeds, Leipzig, Lisbon, Liverpool, London, Madrid, Malaga, Manchester, Marseille, Milan, Nantes, Nice, Nuernberg, Oporto, Palma de Mallorca, Palermo, Rome, Strasbourg, Turin, Valencia, Venice, and Warsaw. However, the majority of the possible combinations among those cities are *not* served by Ryanair, so that only a small number of drawn routes were affected.

These steps give the 92 connections between cities underlying our final sample. For each of them, we randomly assigned one of the two cities as departing city for the outbound flight and the other one as the departing city for the return flight. Table 6 lists the resulting cross-section of routes.

B.2 Affiliate groups

Our baseline measure for competition treats airlines that are affiliated through cross-holdings as single competitors. An airline is matched to an affiliate group if a member of that group owns more than 25% of the airline's equity. Out of the airlines observed in our sample, we have identified the following affiliate groups based on this criterion.

- Aegean, Aegean Airlines, Olympic
- Air France, KLM
- Air One, Alitalia, Meridiana fly, Wind Jet
- British Airways, Iberia, Vueling Airlines
- Air Dolomiti, Austrian Airlines, bmi, Brussels Airlines, Condor, Germanwings, Lufthansa, SunExpress, Swiss International Air Lines
- Blue 1, Cimber Sterling, Norwegian Air Shuttle, SAS, Spanair
- airberlin, Niki
- LAN Airlines, TAM Airlines, TAM Brazilian Airlines
- Singapore Airlines, Virgin Atlantic

Table 6: List of routes

Origin	Destination	Origin	Destination	Origin	Destination
Aberdeen	Manchester	London	Bordeaux	Paris	Dublin
Amsterdam	Barcelona	London	Frankfurt	Paris	Hamburg
Amsterdam	Zurich	London	Hannover	Paris	London
Athens	Budapest	London	Prague	Paris	Madrid
Athens	London	London	Sofia	Paris	Marseille
Barcelona	Lyon	London	Zurich	Paris	Prague
Belgrade	Vienna	Liverpool	Amsterdam	Paris	Stockholm
Berlin	Helsinki	Lyon	Madrid	Paris	Turin
Berlin	Vienna	Madrid	Barcelona	Paris	Valencia
Bilbao	Paris	Madrid	Copenhagen	Paris	Warsaw
Bologna	Madrid	Madrid	Lisbon	Palermo	Turin
Bordeaux	Madrid	Madrid	Milan	Prague	Helsinki
Bordeaux	Nantes	Madrid	Stockholm	Prague	Milan
Brussels	Leeds	Madrid	Valencia	Prague	Rome
Brussels	London	Madrid	Zurich	Rome	Nice
Bucharest	Milan	Malaga	Madrid	Rome	Vienna
Budapest	Munich	Milan	Copenhagen	Stockholm	Berlin
Copenhagen	Geneva	Milan	Duesseldorf	Stockholm	Duesseldorf
Copenhagen	Helsinki	Milan	Frankfurt	Stockholm	Oslo
Duesseldorf	Athens	Milan	Lyon	Stuttgart	Milan
Edinburgh	Manchester	Milan	Paris	Strasbourg	Paris
Frankfurt	Innsbruck	Moscow	Budapest	Toulouse	Brussels
Frankfurt	Istanbul	Munich	Athens	Toulouse	Paris
Frankfurt	Madrid	Munich	Madrid	Vienna	Amsterdam
Frankfurt	Moscow	Munich	Paris	Vienna	Barcelona
Frankfurt	Paris	Munich	Vienna	Vienna	Frankfurt
Frankfurt	Toulouse	Naples	Milan	Vienna	Lyon
Hamburg	Warsow	Nice	Brussels	Vienna	Paris
Hannover	Amsterdam	Nuremberg	Amsterdam	Zurich	Frankfurt
Leipzig	Munich	Oporto	Paris	Zurich	Mallorca
Lisbon	Amsterdam	Paris	Copenhagen		

Notes: Prices are recorded for 41 distinct travel dates for each route. In 10 instances, we did not find any of our roundtrip combinations offered; 7 of them missing on the route Brussels–Leeds; the remaining 3 markets are missing on the routes Bordeaux–Madrid, Moscow–Budapest and Stockholm–Berlin.

- Air Seychelles, Etihad Airways
- Aeroflot-Russian Airlines, Malev Hungarian Airlines

C Robustness specifications

C.1 Alternative competition measures

This appendix contains some robustness analysis with respect to our baseline approach to measure competition.

First, we consider two variations of our baseline approach to identify the number of competitors. Specifically, in our baseline measure, we treat codesharing airlines as competitors. Accordingly, if the same physical connection is marketed under different flight numbers that correspond to different (non-affiliated) airlines, this increases our measure of competition. The reasoning behind this choice is that in so-called “block space” codeshare agreements, each of the codesharing partner still controls a distinct, *ex ante* fixed amount of seats. In practice, by the pricing agreements between the carrier operating a service and the codesharing partner, the codesharer is usually granted considerable freedom to set prices independently.³⁴ In line with that, prices in our data differ indeed substantially across different codesharers.³⁵

To evaluate whether implicit pricing agreements between codesharing airlines systematically affect our findings, we consider N_{ijt}^{csh} as a first alternative, which defines competition as the number of airlines that *operate* their own services on a particular market.³⁶ Similarly, we also consider N_{ijt}^{allie} , which counts the number of competing airline alliances, for which a similar concern might be raised.³⁷ Finally, we also use

³⁴See, e.g., the report by the European Commission, “Competition impact of airline code-share agreements: Final report” (2007), available on the EC Website (last checked: October 2012).

³⁵In our data set, 33.5 percent of the same physical roundtrip combinations ij for a given travel date t and a given date of the ticket offer d are offered by a codesharing operator. Among those observations, the within- $ijtd$ standard deviation across codesharers (i.e., the standard deviation among tickets sold at the same day for the same physical flight by different codesharers) is on average 137.45 Euros. While this large value is driven by a strongly right-skewed distribution, a median standard deviation of 21.73 Euros and an average standard deviation among the first 90 percentiles of 75.09 Euros, suggest that there is indeed significant leeway for independent pricing among codesharing partners.

³⁶To be consistent with this approach, we also pool all physically identical roundtrips into a single observation, where at each date the pooled roundtrip is assigned the lowest price offered by any of the codesharing partners.

³⁷Specifically, N_{ijt}^{allie} treats all airlines within “Star Alliance”, “Sky Team” and “One World” as

the Herfindahl index, which is a common alternative used in the literature.³⁸ For each of these measures, we consider both a log and a level variant.³⁹

Table 7 reports the results to the following estimation:

$$\ln(\text{Price}_{ijt}) = (\alpha + \beta \text{Daysleft}_{td}) \times \text{Comp}_{ijt} + \lambda_i + \mu_j + \nu_t + \xi_d + \epsilon_{ijt}, \quad (8)$$

where Comp_{ijt} is a stand-in for the eight competition measures that we consider, and where λ_i , μ_j , ν_t , and ξ_d are treated as fixed-effects.

The first thing to note, as can be seen in Panel (b), is that all competition measures yield the same qualitative conclusions as our baseline setup. In order to assess the quantitative implications across the variety of competition measures, we compute the linear combination of the baseline slope (Panel a) with the competition effects for the minimum and maximum realization of competition for each of the measures. The implied range of intertemporal slopes for each of our measures is reported in Panel (c). Here it can be seen that also quantitatively, the different measures yield very similar conclusions: While the different measures find a daily price increase ranging from 1.27 percent to 1.39 percent for the least competitive routes in our sample, the daily price increase estimated for the most competitive routes ranges from 0.87 percent to 1.02 percent.

C.2 Alternative customer heterogeneity measures

In this appendix, we consider some alternatives to the customer heterogeneity measures considered in the main body of the paper.

As a first alternative, we use two measures aimed at capturing variations in the customer heterogeneity across time. Specifically, we use dummy variables (Weekend_t and Holidays_t) to distinguish weekend trips (departing on Friday and returning on Sunday) from trips that depart on Monday and return on Thursday, and to identify roundtrips over the Christmas and New Years holidays.

single competitors. Airlines not belonging to any of these alliances are counted as independent competitors according to our baseline competition measure.

³⁸Since a Herfindahl index of one indexes the highest concentration of market power, we use $-H_{ijt}$ to make it qualitatively comparable with the other measures.

³⁹To facilitate the comparison across the different measures which share different supports (the Herfindahl index is defined on the real line, whereas N_{ijt} , N_{ijt}^{csh} , and N_{ijt}^{allie} are natural numbers with different ranges), we adopt linear specifications for each of the competition measures in our estimation.

Table 7: Alternative competition measures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>(a) Baseline Slope</i>								
$Daysle ft_{itd}$	-1.37 (0.09)	-1.32 (0.09)	-1.39 (0.09)	-1.29 (0.09)	-1.36 (0.09)	-1.29 (0.09)	-0.89 (0.09)	-1.38 (0.09)
<i>(b) Competition Effects</i>								
N_{ijt}	0.08 (0.01)							
$\ln N_{ijt}$		0.19 (0.02)						
N_{ijt}^{csh}			0.11 (0.02)					
$\ln N_{ijt}^{csh}$				0.20 (0.03)				
N_{ijt}^{allie}					0.09 (0.01)			
$\ln N_{ijt}^{allie}$						0.18 (0.03)		
$-H_{ijt}$							0.50 (0.05)	
$-\ln H_{ijt}$								0.31 (0.03)
<i>(c) Implied Overall Slope</i>								
Least competitive routes	-1.29	-1.32	-1.28	-1.29	-1.27	-1.29	-1.39	-1.38
Most competitive routes	-0.87	-0.98	-0.94	-1.02	-0.95	-1.02	-1.00	-0.91
Observations	1 417 635	1 417 635	1 417 635	1 417 635	1 417 635	1 417 635	1 417 635	1 417 635
R-squared (adj.)	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58

Notes: The dependent variable is $\ln(Price_{ijt,d})$. Unreported but included in the estimations are levels of the competition measures and fixed effects λ_i , μ_j , ν_t and ξ_d . Clustered standard errors (at the market level) are reported in parentheses. All reported coefficients are significant at the 0.1 percent level and are multiplied by 100. Overall slopes are linear combinations of the baseline slope and competition effects evaluated at the minimum and maximum value of the corresponding competition measure. E.g., in (4), N_{ijt}^{csh} ranges from 1 to 4, implying an overall slope for the most competitive route of $\ln 4 \times 0.1974 - 1.2902$, whereas in (7) the Herfindahl index H_{ijt} is ranging from 1 to 0.2223, implying an overall slope for the least competitive route of $-1 \times 0.5034 - 0.8904$.

The next two alternatives are based on an idea by Brueckner, Dyer and Spiller (1992) who argue that a good proxy for the share of leisure travelers on a route are temperature differentials between the destination and origin.⁴⁰ Accordingly, we define $\Delta Temp_{ijt}$ as the temperature differentials between destination and origin during the month of travel. In the same spirit, we also use $\Delta Prec_{ijt}$ to measure precipitation differentials between destination and origin during the month of travel.

Finally, we use regional GDP data to measure for business activity.⁴¹ Accordingly, GDP_{ij}^o and GDP_{ij}^d are defined as log GDP at the origin and destination, respectively.

To give an overview over these measures, Figure 5 plots their means (marked by circled dots), standard deviation bands (marked by boxes), and supports (marked by lines) conditional on the number of competitors. Again the correlation across competition levels is typically small relative to the standard deviation and supports, although somewhat larger than for our two baseline measures.

To investigate how changes in customer heterogeneity affect the sensitivity of the intertemporal slope to competition, we use the same empirical model as in Section 5:

$$\ln(Price_{ijt}) = (\alpha + \beta Daysleft_{td}) \times X_{ijt} + \lambda_i + \mu_j + \nu_t + \xi_d + \epsilon_{ijt}, \quad (9)$$

with

$$X_{ijt} = (1, N_{ijt}, Het_{ijt}, N_{ijt} \times Het_{ijt}),$$

where N_{ijt} is the number of competitors, Het_{ijt} is the corresponding customer heterogeneity measure, and λ_i , μ_j , ν_t and ξ_d are treated as fixed effects. Again, we normalize Het_{ijt} to have zero mean and unit variance. Here β is the relevant vector of “slope” coefficients that define the intertemporal slope. Table 8 contains the estimated coefficients and standard errors (clustered at the market level). For convenience, Columns 1 and 2 reprint the results for our baseline measures.

As before, we look at the impact of Het_{ijt} on the slope in monopoly markets to determine whether it is likely to increase or decrease customer heterogeneity.

⁴⁰See also Stavins (2001) and Goolsbee and Syverson (2008), who have used similar temperature-based proxies to measure for tourist intensities.

⁴¹We use GDP data provided by the European commission at the NUTS-2 level. See also Footnote 16.

Table 8: Alternative customer heterogeneity measures

Customer Heterogeneity Measure:	(1) $Duration_{ij}$	(2) $Intensity_{ij}$	(3) $Weekend_t$	(4) $Holidays_t$	(5) $\Delta Temp_{ijt}$	(6) $\Delta Prec_{ijt}$	(7) GDP_{ij}^o	(8) GDP_{ij}^d
<i>(a) Estimated Slope Coefficients</i>								
$Daysle.ft_{td}$	-1.41 (0.10)	-1.37 (0.10)	-1.00 (0.09)	-1.33 (0.09)	-1.36 (0.09)	-1.34 (0.09)	-1.43 (0.09)	-1.38 (0.10)
$Daysle.ft_{td} \times N_{ijt}$	0.07 (0.01)	0.08 (0.01)	0.08 (0.01)	0.08 (0.01)	0.08 (0.01)	0.07 (0.01)	0.09 (0.01)	0.07 (0.01)
$Daysle.ft_{td} \times Het_{ijt}$	-0.13 (0.02)	-0.16 (0.02)	0.25 (0.02)	0.09 (0.02)	0.02 (0.02)	-0.06 (0.02)	0.11 (0.03)	-0.07 (0.03)
$Daysle.ft_{td} \times N_{ijt} \times Het_{ijt}$	0.06 (0.01)	0.08 (0.01)	-0.06 (0.01)	-0.02 (0.01)	-0.02 (0.01)	0.04 (0.01)	-0.03 (0.01)	0.02 (0.01)
<i>(b) Implied Sensitivity to Competition and Overall Slope</i>								
Low heterogeneity markets								
sensitivity	0.01	-0.01	0.02	0.06	0.10	0.03	0.06	0.05
monopolies	-1.27	-1.22	-0.73	-1.18	-1.28	-1.25	-1.26	-1.25
competitive	-1.22	-1.24	-0.64	-0.88	-0.80	-1.10	-0.96	-0.99
High heterogeneity markets								
sensitivity	0.13	0.16	0.14	0.10	0.05	0.10	0.13	0.09
monopolies	-1.42	-1.38	-1.11	-1.31	-1.28	-1.29	-1.41	-1.35
competitive	-0.77	-0.60	-0.39	-0.82	-1.01	-0.79	-0.78	-0.88
Classification of Het_{ijt}	+	+	-	-	+ [†]	+	-	+
Observations	1 255 125	1 286 687	1 417 635	1 417 635	1 417 635	1 417 635	1 377 958	1 286 687
R-squared (adj.)	0.57	0.57	0.58	0.58	0.58	0.58	0.58	0.57

Notes: The dependent variable is $\ln(Pric_{eijt})$. Unreported but included in the estimations are levels of $Daysle.ft_{td}$, N_{ijt} , Het_{ijt} and $N_{ijt} \times Het_{ijt}$, and fixed effects λ_i , μ_j , ν_t and ξ_d . The variable used for Het_{ijt} is indicated at the top of each column. Het_{ijt} is transformed to have zero mean and unit variance. Clustered standard errors (at the market level) are reported in parentheses. Reported coefficients and standard errors are multiplied by 100. All reported coefficients are significant at the 1 percent level, except the slope coefficients on Het_{ijt} in Columns 5 and 6 (in Column 6, it is significant at the 5 percent level), and on $N_{ijt} \times Het_{ijt}$ in Columns 4 and 8 (which are significant at the 5 percent level). Increases in Het_{ijt} that steepen the intertemporal slope in monopolies are classified as increasing heterogeneity (indicated at bottom of table). [†] indicated that the classification is not significant at the 5 percent level. Low and high heterogeneity markets in Panel (b) are defined as markets with $Het_{ijt} = \pm 1$ standard deviation relative to the mean, with the direction determined by the classification of Het_{ijt} . Competitive markets are defined as markets with six competitors. Number of observations depends on availability of heterogeneity data.

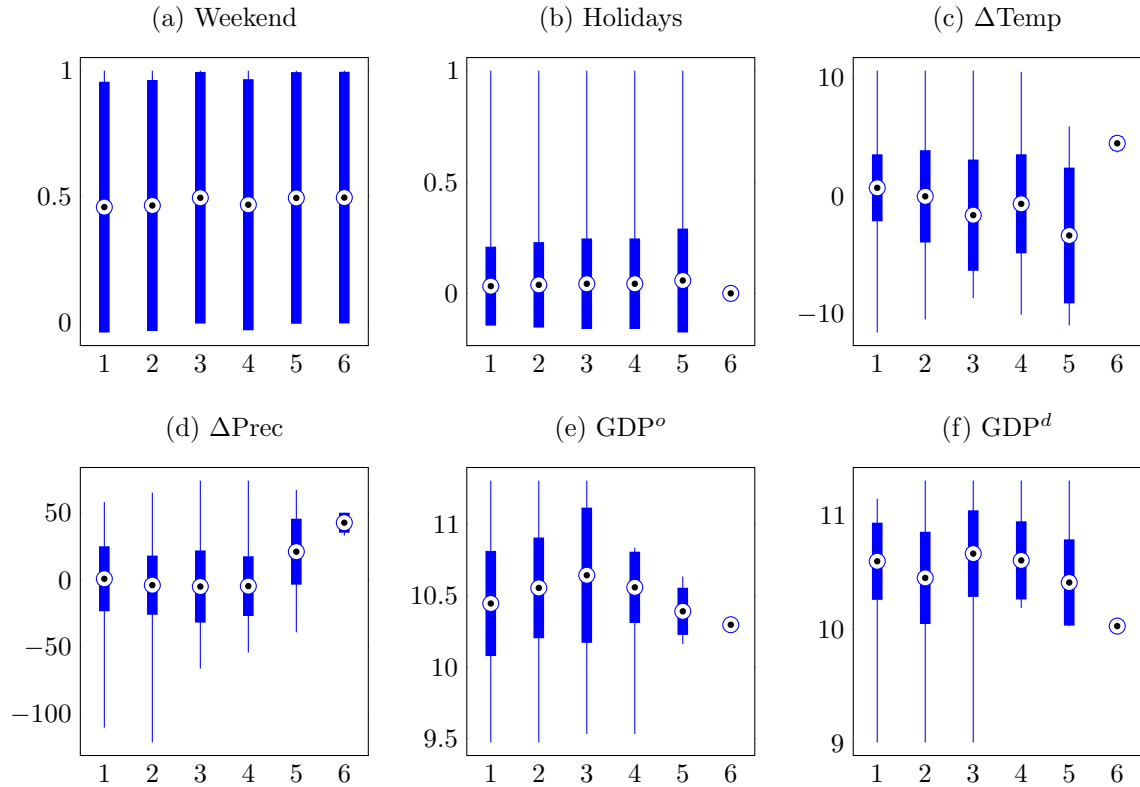


Figure 5: Summary of customer heterogeneity measures by competition.
Notes: Means and ± 1 standard deviation bands are indicated by the circled dots and boxes. Supports are indicated by the extending lines. Competition categories are indicated on the x-axis.

Accordingly, we interpret Het_{ijt} as *increasing* in the customer heterogeneity if

$$\left. \frac{\partial^2 \ln(\text{Price}_{ijt})}{\partial \text{Daysleft}_{td} \partial Het_{ijt}} \right|_{N_{ijt}=1} = \beta_{Het} + \beta_{N \times Het} < 0,$$

and vice versa. Applying this criterion to the estimated models, we classify increases in $\Delta Prec_{ijt}$ and GDP_{ij}^d as increasing customer heterogeneity, and classify increases in $Weekend_t$, $Holidays_t$ and GDP_{ij}^o as decreasing customer heterogeneity. For $\Delta Temp_{ijt}$, its impact on the intertemporal slope is not statistically different from zero, both economically and statistically (at the 5 percent level). With this qualification in mind, increases in $\Delta Temp_{ijt}$ are increasing customer heterogeneity according to our classification approach. For convenience, these classifications are listed at the bottom of Table 8.

Recall that for the sensitivity to increase in more heterogeneous markets, the

slope coefficient for $N_{ijt} \times Het_{ijt}$ has to go into the same direction as the impact of Het_{ijt} on heterogeneity. That is, if we have associated Het_{ijt} with an increase in heterogeneity, then the slope coefficient for $N_{ijt} \times Het_{ijt}$ has to be positive in order for the intertemporal slope to be more sensitive to increases in competition in more heterogeneous markets, and vice versa.

Comparing the above classifications with the slope coefficients on $N_{ijt} \times Het_{ijt}$, these estimations confirm the conclusions based on hotel bookings data for all measures except $\Delta Temp_{ijt}$ (for which we could not reject the possibility of a false classification at the 5 percent level). Looking again at the sensitivity to competition for Het_{ijt} at one standard deviation above and below its mean,

$$\frac{\partial^2 \ln(Price)}{\partial Daysleft_{td} \partial N_{ijt}} \Big|_{Het_{ijt}=\pm 1} = \beta_N \pm \beta_{N \times Het}.$$

we see that all measures with the exception of $\Delta Temp_{ijt}$ yield similar patterns (but to a lesser extent) to the ones seen for our baseline measures. The difference is likely to result from a somewhat looser correlation between these alternative measures and customer heterogeneity, so that evaluation Het_{ijt} at one standard deviation above and below its mean is likely to map into an actual customer heterogeneity evaluated substantially within its one standard deviation circle.

C.3 Non-monotonic impact of customer heterogeneity

In this section, we generalize our approach in the main body of the paper to evaluate how the sensitivity of the intertemporal slope to competition changes with customer heterogeneity in a potential non-monotonic fashion.

We use the following quadratic specification in Het_{ijt} to allow for such non-monotonicities:

$$\ln(Price_{ijtd}) = (\alpha + \beta Daysleft_{td}) \times X_{ijt} + \lambda_i + \mu_j + \nu_t + \xi_d + \epsilon_{ijtd}, \quad (10)$$

with

$$X_{ijt} = (1, N_{ijt}, Het_{ijt}, Het_{ijt}^2, N_{ijt} \times Het_{ijt}, N_{ijt} \times Het_{ijt}^2),$$

where N_{ijt} is the number of competitors, Het_{ijt} is the corresponding customer heterogeneity measure, and λ_i , μ_j , ν_t and ξ_d are treated as fixed effects. To ensure that

computing the quadratic terms is a positive monotonic transformation, we normalize Het_{ijt} to have positive support.

To identify whether a certain change in Het_{ijt} increases or decreases heterogeneity, we generalize our previous approach to look at the impact of Het_{ijt} on the intertemporal slope in monopoly markets. In particular, we interpret a change in Het_{ijt} as an increase in customer heterogeneity if

$$\frac{\partial^2 \ln(Price_{ijt_d})}{\partial Daysleft_{td} \partial Het_{ijt}} \Big|_{N_{ijt}=1} = \beta_{Het} + \beta_{N \times Het} + 2(\beta_{Het^2} + \beta_{N \times Het^2}) Het_{ijt} < 0,$$

and vice versa. Note that through the quadratic term, the classification may now change with Het_{ijt} . For example, we could have that the heterogeneity is increasing in the tourist intensity if the tourist intensity is absolute low, but is decreasing once the market is satiated with tourists.

Based on this criterion, we do not find any evidence for a non-monotonic impact on customer heterogeneity. Both $Duration_{ijt}$ and $Intensity_{ij}$ have a positive impact on heterogeneity throughout their complete support, motivating the non-monotonic treatment in the main body of the paper.

Figure 6 plots the estimated sensitivity of the intertemporal slope to competition as a function of Het_{ijt} , formally given by the cross-derivative of the identified model in $Daysleft_{td}$ and N_{ijt} :

$$\frac{\partial^2 \ln(Price_{ijt_d})}{\partial Daysleft_{td} \partial N_{ijt}} = \beta_N + \beta_{N \times Het} Het_{ijt} + \beta_{N \times Het^2} Het_{ijt}^2.$$

Not surprising given the monotonic impact of Het_{ijt} , the sensitivity of the slope to competition is literally “in line” with the linear setup in the main body, yielding a sensitivity that ranges from zero at the 5th percentile of Het_{ijt} (largely homogenous markets by our classification) to somewhat above 0.15 and 0.20, respectively, at the 95th percentile (substantially heterogeneous markets by our classification).

C.4 Competition-specific impact of customer heterogeneity

This appendix section generalizes the analysis in the main body of the paper to allow for a competition-specific impact of customer heterogeneity.

To investigate how changes in customer heterogeneity affect the sensitivity of the

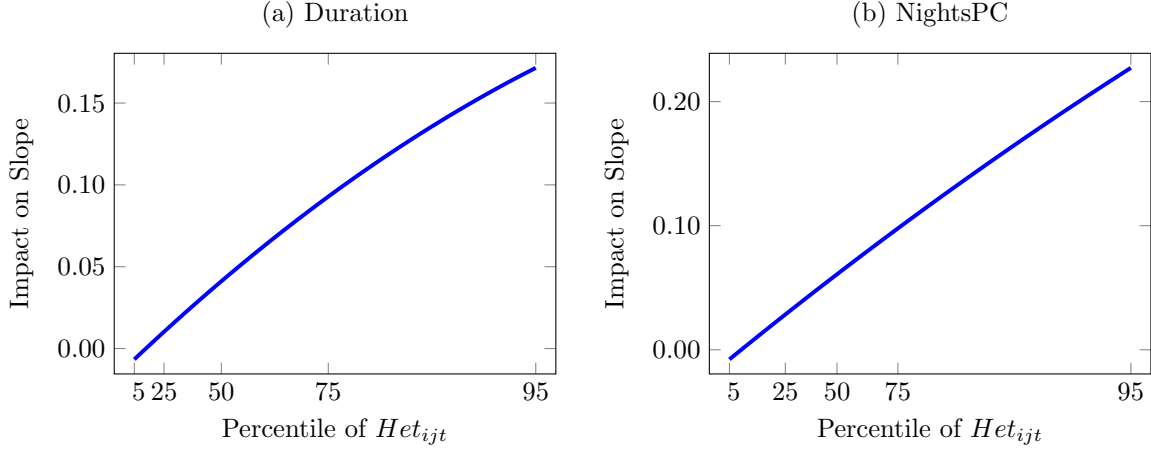


Figure 6: Sensitivity of intertemporal slope to competition.

Notes: Plots display the impact of N_{ijt} on the intertemporal slope as a function of Het_{ijt} , resulting from a quadratic specification in Het_{ijt} . On the x-axes, Het_{ijt} is scaled proportional to its own scale with labels displaying the corresponding percentiles. The y-axes display the cross-derivative of the estimated model in N_{ijt} and $Daysleft_{td}$. Het_{ijt} are properly normalized to have a positive support. The estimations control for levels of N_{ijt} , Het_{ijt} , Het_{ijt}^2 , $N_{ijt} \times Het_{ijt}$, $N_{ijt} \times Het_{ijt}^2$; the same variables interacted with $Daysleft_{td}$; and fixed effects λ_i , μ_j , ν_t and ξ_d .

intertemporal slope to competition, we estimate the following model:

$$\ln(Price_{ijt}) = (\alpha + \beta Daysleft_{td}) \times X_{ijt} + \lambda_i + \mu_j + \nu_t + \xi_d + \epsilon_{ijt}, \quad (11)$$

with

$$X_{ijt} = (Comp_{ijt}, Comp_{ijt} \times Het_{ijt}),$$

where Het_{ijt} is the corresponding customer heterogeneity measure, $Comp_{ijt} \times Het_{ijt}$ is a vector of interaction effects between Het_{ijt} and all competition dummies⁴², and λ_i , μ_j , ν_t and ξ_d are treated as fixed effects. To facilitate the discussion below, we normalize Het_{ijt} to have zero mean and unit variance within the corresponding competition category when interacting with $Comp_{ijt}$. Here β is the relevant vector of “slope” coefficients that define the intertemporal slope. The impact of Het_{ijt} for $Comp_{ijt} = 6$ is not identifiable due to zero variation (see Figure 4). We therefore pool 5 and 6 competitors to a “ $Comp_{ijt} = 5+$ ” category. Table 9 contains the estimated coefficients and standard errors (clustered at the market level).

From Panel (a) it can be seen that the non-interacted slope coefficients are again

⁴²Since $Comp_{ijt}$ covers all competition categories there is no need to include Het_{ijt} by itself.

Table 9: Customer heterogeneity and intertemporal price dispersion by market competitiveness

Customer Heterogeneity Measure:	(1) <i>Duration_{ij}</i>	(2) <i>Intensity_{ij}</i>
<i>(a) Estimated Slope Coefficients</i>		
$(Comp_{ijt} = 1) \times Daysleft_{td}$	-1.36 (0.10)	-1.30 (0.10)
$(Comp_{ijt} = 2) \times Daysleft_{td}$	-1.26 (0.10)	-1.22 (0.10)
$(Comp_{ijt} = 3) \times Daysleft_{td}$	-1.22 (0.10)	-1.18 (0.10)
$(Comp_{ijt} = 4) \times Daysleft_{td}$	-1.12 (0.10)	-1.08 (0.10)
$(Comp_{ijt} = 5+) \times Daysleft_{td}$	-1.02 (0.10)	-0.98 (0.10)
$(Comp_{ijt} = 1) \times Het_{ijt} \times Daysleft_{td}$	-0.10 (0.02)	-0.11 (0.02)
$(Comp_{ijt} = 2) \times Het_{ijt} \times Daysleft_{td}$	-0.00 (0.01)	-0.02 (0.02)
$(Comp_{ijt} = 3) \times Het_{ijt} \times Daysleft_{td}$	0.08 (0.02)	0.05 (0.01)
$(Comp_{ijt} = 4) \times Het_{ijt} \times Daysleft_{td}$	0.11 (0.03)	0.18 (0.02)
$(Comp_{ijt} = 5+) \times Het_{ijt} \times Daysleft_{td}$	0.17 (0.04)	0.17 (0.04)
<i>(b) Sensitivity and Overall Slope</i>		
Low heterogeneity markets		
average sensitivity to competition	0.02	0.01
implied monopoly slope	-1.25	-1.19
implied competitive slope ($N_{ijt} = 5+$)	-1.19	-1.15
High heterogeneity markets		
average sensitivity to competition	0.15	0.15
implied monopoly slope	-1.46	-1.41
implied competitive slope ($N_{ijt} = 5+$)	-0.84	-0.81
Observations	1 255 125	1 286 687
R-squared (adj.)	0.57	0.57

Notes: The dependent variable is $\ln(Price_{ijt_d})$. Unreported but included in the estimations are levels of $Comp_{ijt}$, $Comp_{ijt} \times Het_{ijt}$, and fixed effects λ_i , μ_j , ν_t and ξ_d . The variable used for Het_{ijt} is indicated at the top of each column. Het_{ijt} is transformed to have zero mean and unit variance *within* the corresponding competition category when interacted with competition. Clustered standard errors (at the market level) are reported in parentheses. Reported coefficients and standard errors are multiplied by 100. All coefficients are significant at the 0.1 percent level, except for $(Comp_{ijt} = 2) \times Het_{ijt} \times Daysleft_{td}$. Low and high heterogeneity markets in Panel (b) are defined as markets with $Het_{ijt} = \pm 1$ standard deviation relative to the mean in the corresponding competition category. Number of observations depends on availability of heterogeneity data.

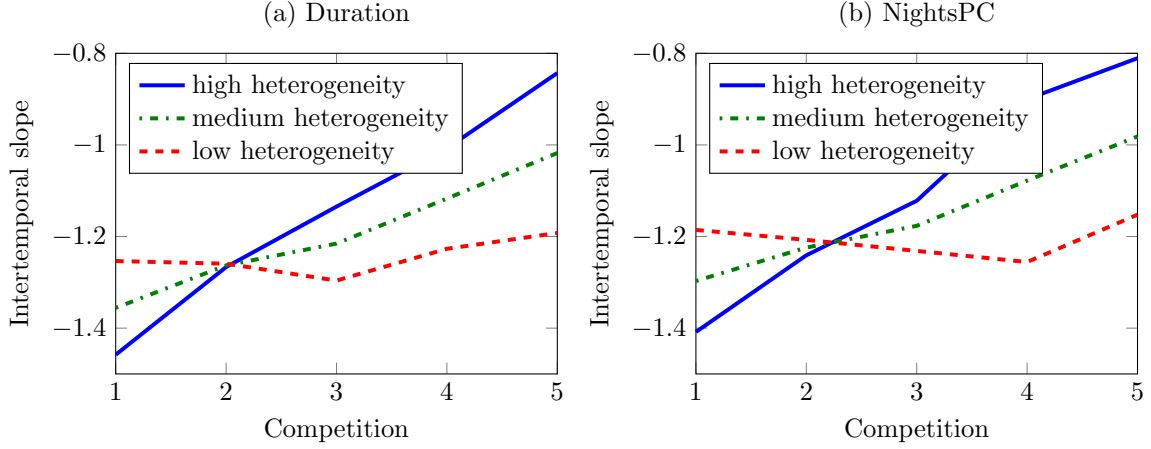


Figure 7: Empirical relationships between intertemporal slope and competition for different degrees of customer heterogeneity.

Notes: The intertemporal slopes are denominated in daily price changes in percent. The medium heterogeneity lines are evaluated at the mean value (within-competition) of the corresponding tourist measure. The high and low heterogeneity lines are evaluated at the mean ± 1 standard deviation. The tourist measure underlying panel (a) is $Duration_{ij}$; the measure underlying panel (b) is $Intensity_{ij}$. All computations are based on the regressions reported in Table 9.

similar to our baseline model in Section 4, likely because there is little correlation between Het_{ijt} and $Comp_{ijt}$. Accordingly, prices are again unambiguously increasing as the scheduled travel date approaches, but at an intertemporal slope that “flattens” in more competitive environments.

To evaluate the hypothesis that the intertemporal slope is more sensitive to competition in more heterogeneous markets, we again need to take a stand on whether our tourist intensity measures increase or decrease the customer heterogeneity in a given market. As before, we interpret changes in Het_{ijt} that steepen the intertemporal slope on monopoly routes as *increases* in the customer heterogeneity. Given model (11), this is true if

$$\left. \frac{\partial^2 \ln(Price_{ijt})}{\partial Daysleft_{td} \partial Het_{ijt}} \right|_{N_{ijt}=1} = \beta_{(Comp=1) \times Het} < 0,$$

which—as in the main body of the paper—is the case for both our measures.

To give an overview over how changes in competition affect the intertemporal slope, Figure 7 plots the intertemporal slope as a function of the number of competitors at different values of Het_{ijt} . Similar to the thought experiment conducted in the main body of the paper, we plot the intertemporal slope for “low” and “high” heterogeneity markets (evaluated at one standard deviation below and above the competition-specific

mean of Het_{ijt}), as well as for a “medium” heterogeneity market evaluated at the mean.

While the linear model in the main body assumes that the intertemporal slope is linear in competition (with the derivative defining the “sensitivity” to competition), the competition-specific setup considered here allows for an arbitrary relation. Reassuring our findings in the main body, we however find that the sensitivity is nearly constant conditional on Het_{ijt} . As before, the intertemporal slope is highly sensitive to competition in high heterogeneity markets, as is predicted by theories of competitive price discrimination. In contrast, reducing the customer heterogeneity in a given market to one standard deviation below its mean, makes the intertemporal slope largely invariant to competition, in line with the idea that in homogeneous markets airlines do not price-discriminate to begin with so that they are also not restrict in their ability to do so by competition.

To reconcile these findings quantitatively with the linear model in the main body, Panel (b) of Table 9 also reports the “average” sensitivity to competition, amounting to the average change in the intertemporal slope displayed in Figure 7:

$$\frac{1}{supp(N)} \int_N \frac{\partial^2 \ln(Price_{ijt})}{\partial Daysleft_{ijt} \partial N_{ijt}} \Big|_{Het_{ijt}=\pm 1} \approx \frac{1}{4} \sum_{N=2}^5 \left\{ \left[\beta_{Comp=N} - \beta_{Comp=N-1} \right] \pm \left[\beta_{(Comp=N) \times Het} - \beta_{(Comp=N-1) \times Het} \right] \right\}.$$

Not surprising given the nearly linear relation in Figure 7, the implied sensitivities are numerically consistent with our linear model, motivating the simplified exposition in the main body of the paper.

D Additional figures

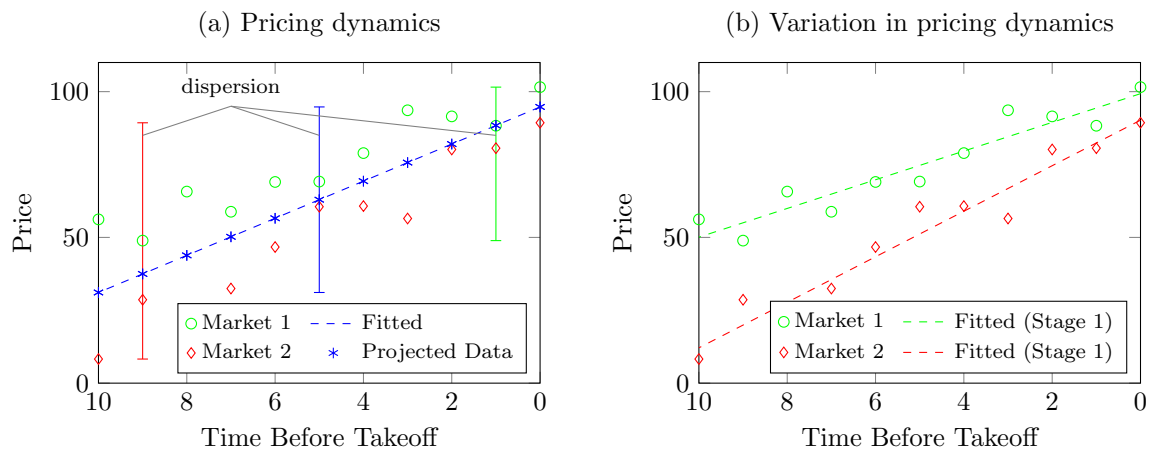


Figure 8: Unsystematic pricing dynamics versus unsystematic variations in the intertemporal slope. *Notes:* The figure displays fictitious price data in 2 markets. Panel (a) displays how projected data based on a linear fit explains most of the dispersion (86 percent of the average Gini of markets 1 and 2)—reflecting that systematically increasing prices are the main source behind the observed dispersion. Panel (b) illustrates how at the same time, the fitted model fails to explain any of the variation in the market-specific slopes. The left panel is akin to the exercise conducted in Section 7.1. Section 7.2 explores, in contrast, how much of the variation in slopes depicted in Panel (b) is explainable by variations in customer heterogeneity.

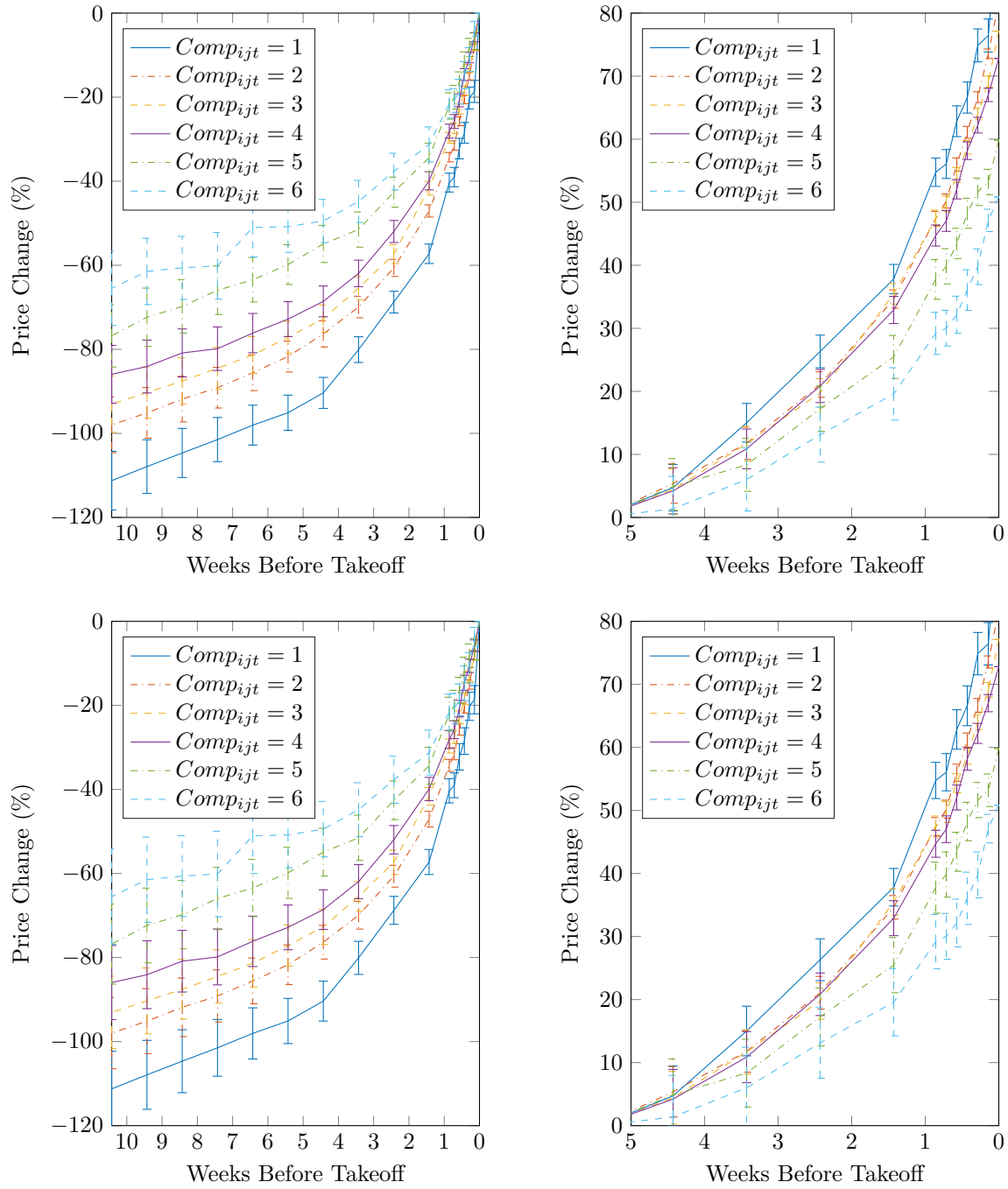


Figure 9: Nonlinear estimation of pricing dynamics with $p = 0.80$ (top panels) and $p = 0.90$ (bottom panels) confidence bands.

Notes: Coefficients are multiplied by 100 and are normalized relative to the prices 1 day (left panel) and 38 days (right panel) before takeoff. To increase readability, the right panels truncate both the range and the support of the data. The relative price changes prior to 5 weeks before departure are *not* statistically different from each other. For a complete picture and further details see Figure 3 and the description in the main text.