Costs, Discount Factors and Coordination in the Airline Industry*

Federico Ciliberto† Alon Eizenberg‡ Max Schnidman§

February 15, 2025

Abstract

We develop an econometric method to set-identify the degree of potential deviation from competitive (Nash) behavior across airline markets. The approach leverages the presence of a common unobserved factor, the rental rate of capital, in marginal costs and discount factors. We embed this dual constraint into a structural model of repeated interaction. In addition to standard price and quantity data, we use a novel dataset on aircraft market values to pin down key components of the cost structure, namely, the cost of capital and depreciation. Applying the methodology to the US domestic airline industry in 2018-2019, we cannot reject full coordination in about two-thirds of markets with multiple legacy carriers. In addition, our tests more frequently reject the hypothesis of coordination in markets featuring low-cost carriers, with Southwest Airlines' presence showing the strongest negative correlation with our measures of potential coordination.

*First Draft. Preliminary and incomplete. Please do not cite or distribute without permission.

†Department of Economics, University of Virginia, DIW and CEPR. ciliberto@virginia.edu

‡Department of Economics, Hebrew University of Jerusalem, and CEPR. alon.eizenberg@mail.huji.ac.il

§Department of Economics, University of Virginia. mes7jw@virginia.edu

1 Introduction

Measuring the intensity of competition and understanding its determinants are crucial areas of study in economics. While competitive conduct is often linked to observable or estimable market characteristics such as the number of competitors or cost and demand structures, the strength of competition can vary significantly even when these factors are similar. It is therefore challenging to identify which markets are truly competitive and to understand the underlying reasons.

Our paper addresses this challenge by proposing an empirical framework that *set-identifies* the degree of competitive conduct across a range of airline markets. We then analyze how various market characteristics correlate with these competitive measures to test hypotheses about what drives the intensity of competition. This approach integrates several key strands of the industrial organization literature, offering a comprehensive method for studying market dynamics.

We study how firms interact over time within a repeated game framework with a specific focus on the airline industry. In such games, the discount factor plays a critical role in sustaining supra-competitive equilibria. However, the discount factor is not directly observable. Furthermore, the presence of multiple equilibria suggests that even when the discount factor is large, i.e., firms are more patient, competitive equilibria may still prevail.

Our approach bridges the perspectives of Bresnahan and Schmalensee, reflected in their respective chapters in the 1989 Handbook of Industrial Organization.[1] Bresnahan highlights that marginal costs are not directly observable, while Schmalensee attributes this unobservability specifically to the rental rate of capital.

Our contribution lies in exploiting a dual constraint: *the rental rate of capital must appear in both the discount factor (influencing collusive behavior) and the marginal cost.* By collecting a novel dataset on legacy carriers' capital prices alongside traditional price and quantity data, we estimate a structural model of demand and supply that imposes this constraint. This approach enables set identification of both the rental rate of capital and market-specific bounds on the degree of coordination consistent with the data.

Our empirical strategy embeds competition within a repeated game framework. To measure the degree of coordination, we introduce a conduct parameter $\kappa \in [0, 1]$ that indexes departure from the Nash benchmark. While $\kappa = 0$ reflects the static Nash outcome and $\kappa = 1$ represents full coordination, intermediate values capture partial coordination. While such conduct parameters are often used in static games, we follow Eizenberg et al. (2023) and apply κ to measure coordination on the equilibrium path of a repeated game.

Our set identification of κ builds on a simple but powerful insight. Each value of κ generates

[1] Bresnahan (1989), Schmalensee (1989).

two key objects in our model: first, a threshold discount factor δ needed to sustain that level of coordination in the repeated game; second, a vector of marginal costs (derived from pricing first-order conditions) that implies a rate of return on capital r. Since the discount factor and rate of return are inherently linked through $\delta = 1/(1+r)$, we can test whether any given κ is consistent with the data. Specifically, if the implied discount factor exceeds the threshold needed to sustain that level of coordination, we cannot reject that value of κ. The set of non-rejected κ values forms our identified set for the degree of coordination in each market.

To derive the implied rate of return on capital from marginal costs, we model airline production using a Leontief technology that combines aircraft (capital), labor, and jet fuel in fixed proportions. The resulting route-level cost function depends on four key components: the return on capital, the cost of capital, depreciation, and the load factor (which measures capacity utilization). We estimate the cost of capital and depreciation using aircraft replacement cost data, and observe load factors directly. This allows us to back out the implied rate of return on capital from our estimated marginal costs.

We apply this methodology to US domestic airline markets (airport pairs) in 2018-2019, focusing on potential coordination among legacy carriers while treating low-cost carriers as Nash players. For each market, our estimated set of feasible coordination levels takes an interval form $(0, \overline{\kappa}]$, where $\overline{\kappa} \leq 1$ represents the highest degree of coordination that our model cannot reject.

In markets with multiple legacy carriers, we cannot reject full coordination ($\kappa = 1$) in about *two-thirds* of cases, suggesting a substantial potential for coordinated behavior. We interpret $\overline{\kappa}$ as an indicator of the scope for coordination in a market and examine how it varies across markets and which factors correlate with higher values of this indicator. Of course, we must note that our inability to reject coordination does not necessarily imply its existenceour results indicate only whether coordination is consistent with the data and our model's assumptions.

Building on these insights, we analyze how $\overline{\kappa}$ varies across markets, revealing patterns that align with theoretical predictions. Most notably, the presence of non-legacy carriers correlates with higher rejection rates of coordination, with Southwest Airlines emerging as the strongest disruptive force. This finding is particularly striking: while Southwest pioneered the low-cost model, it is now an established carrier, and others have adopted more aggressive pricing strategies as mavericks (Alaska, JetBlue) or ultra low-cost carriers (Spirit, Frontier, Sun Country). That Southwest's presence continues to show the strongest negative correlation with coordination among legacy carrierswhether measured by rejection of full coordination or by $\overline{\kappa}$suggests its enduring competitive impact in the routes it serves. While other non-legacy carriers also correlate negatively with coordination levels, their effects are more modest, albeit becoming comparable to Southwest's when examining the continuous measure $\overline{\kappa}$ rather than just full coordination.

We also find that coordination is more frequently rejected in markets where legacy carriers have asymmetric market shares, aligning with theoretical predictions that symmetric firms coordinate more easily. Interestingly, while the identity of non-legacy carriersparticularly Southweststrongly influences coordination patterns, we find no clear relationship between coordination and the presence of specific legacy carriers on a route.

Literature (selected). The theory of oligopolistic coordination in repeated games holds a considerable body of work including seminal contributions by Friedman (1971), Porter (1983), Green and Porter (1984), and Abreu (1986), among others. A smaller yet expanding literature shows how to take such models to data with applications to medical services (Goto and Iizuka (2016)), vitamins (Igami and Sugaya (2022)), beer (Miller et al. (2021)) and fast-moving consumer goods (Eizenberg et al. (2023)).

Coordination in the airline industry has also been studied extensively. Ciliberto and Williams (2014) finds a positive relationship between the degree of coordination, modeled as a conduct parameter in a static pricing game, and the intensity of multimarket contact. Turner (2021) and Kim and Park (2023) find evidence of coordinated effects in the US Airways - American Airlines merger. Additional work includes Ciliberto et al. (2019) and Aryal et al. (2022), the latter examining a mechanism of coordination via public communication.

In light of this evidence, it would appear that the degree of coordination is a nontrivial function of various elements, some of which being "soft" and difficult to quantify and document empirically. Our paper adds to this literature by proposing a method for systematically learning about the likely ease of coordination and attempting to link it to potential determinants.

The remainder of this paper proceeds as follows. Section 2 spells out the repeated game model and its set identification. Section 3 presents the data used in this research. Sections 4 and 5 then describe the empirical implementation and results pertaining to the demand and supply sides of the model, respectively. Section 6 offers concluding remarks.

2 Model

We now describe our structural model, beginning in Section 2.1 with a description of the repeated game among legacy carriers. We model pricing interactions as a repeated game, focusing on carriers' price-setting behavior rather than long-term strategic investments in capacity or hub development. On the equilibrium path, legacy carriers may sustain supra-competitive prices and profits if they are sufficiently patient. Section 2.2 describes our assumptions on demand, production and costs. This structure delivers a partial identification strategy for the degree of coordination in each market, as outlined in Section 2.3.

2.1 The supergame

Consider a cross-section of airline markets $m \in \mathcal{M}$ and single-product carriers operating in each such market indexed by $j \in \mathcal{J}_m$. A subset of the market's carriers $\mathcal{J}_m^L \subset \mathcal{J}_m$ are legacy carriers who engage in a repeated game where firm j's discunt factor is δ_j. The game is played in every quarter, and year-quarters are indexed by t. The remaining firms are Low-Cost Carriers (LCCs) and behave as static Nash players: given any history of the game, and fixing rivals' actions, they set prices to maximize current profits.

In this repeated game, legacy carriers follow grim trigger strategies while low-cost carriers always play Nash-Bertrand. When legacy carriers cooperate, they can maintain prices above competitive levels, and due to strategic complementarity, this raises low-cost carriers Nash-Bertrand prices as well. If any legacy carrier deviates from cooperation, the market permanently reverts to competitive Nash-Bertrand pricing for all firms. We treat fixed costs, entry, and exit as given.[2]

Omitting the time index momentarily, we denote the conduct parameter associated with market m by $\kappa_m \in [0, 1]$. Following Eizenberg et al. (2023) we define the price vector induced by a generic value of κ_m as follows:

Definition 1. *In market m, on the equilibrium path, stage-game prices p^{κ_m} are defined as follows. Each legacy carrier $j \in \mathcal{F}_m^L$ sets the price $p_{j,m}$ by solving:*

$$p_{j,m}^{\kappa_m} = argmax_{p_j} v_j(p_j, p_{-j}^{\kappa_m}) + \kappa_m \sum_{g \in \mathcal{F}_m^L \setminus j} v_g(p_g, p_{-g}^{\kappa_m}),$$

where $v_j(\cdot)$ denotes the variable profit generated by product j given a vector of market prices, and $p_{-j}^{\kappa_m}$ denotes the portion of market m's price vector pertaining to firm j's rivals' prices.

The prices of any LCC carrier, $j \notin \mathcal{F}_m^L$, are given by: $p_j^{\kappa_m} = argmax_{p_j} v_j(p_j, p_{-j}^{\kappa_m})$.

The conduct parameter κ_m measures how far market outcomes deviate from the competitive benchmark. While firms in our model play a repeated game and ultimately receive only their own profits, the resulting prices can be characterized through an "as if" interpretation: the observed prices are mathematically equivalent to those that would arise if, in a static game, each legacy carrier maximized a weighted sum of profits - putting full weight on its own profits, while applying a weight of κ_m to its Legacy rivals' profits. This interpretation helps us quantify the degree of coordination, even though firms do not share in their rivals' profits.

[2]This exogeneity is a feature of much of the theoretical literature and all of the empirical literature on supergames cited in the introduction.

$$v_{jt}(\tilde{p}_j^{\kappa_m}, p_{-j}^{\kappa_m}) + \frac{\delta_j}{1-\delta_j} v_{jt}(p^{NB,m}) \leq \frac{1}{1-\delta_j} v_{jt}(p^{\kappa_m}),\ \forall j \in \mathcal{F}_m^L. \tag{1}$$

Incentive Compatibility Constraints. The Incentive Compatibility Constraint (ICC) compares two scenarios. On the left side is the payoff from deviation: an immediate profit from choosing the optimal deviation price $\tilde{p}_j^{\kappa_m}$ while others maintain equilibrium prices, followed by perpetual Nash-Bertrand profits in all future periods. Such a deviation results in a flow payoff of $v_j(\tilde{p}_j^{\kappa_m}, p_{-j}^{\kappa_m})$. The deviation results in a discounted flow of Nash-Bertrand profits starting in the next period, with these profits maintained at their current-period levels, consistent with the grim-trigger mechanism.[3] The right side shows the discounted stream of profits from continued cooperation. The ICC requires that cooperation profits exceed deviation profits.[4]

The carrier's constraint (1) implies a threshold level of the firm's discount factor which is necessary to support the degree of coordination κ_m in quarter t:

$$\delta_{jt} \geq \frac{v_{jt}(\tilde{p}_j^{\kappa_m}, p_{-j}^{\kappa_m}) - v_j(p^{\kappa_m})}{v_{jt}(\tilde{p}_j^{\kappa_m}, p_{-j}^{\kappa_m}) - v_{jt}(p^{NB,m})} \equiv \underline{\delta}_{jmt}(\kappa_m). \tag{2}$$

The rate of return on capital. The novel aspect of this paper is to connect the above framework with the concept of the return on capital. Firm j's ICC would hold at larger values of κ_m the more patient the firm is, i.e., given higher levels of its discount factor. This discount factor is connected to the firm's implicit rate of return on capital, r_{jt}, as follows:

$$\delta_{jt} = \frac{1}{1+r_{jt}}. \tag{3}$$

If we could observe each legacy carrier's return on capital r_{jt}, we could determine their discount factors δ_{jt}.[5] This would have allowed us to solve equation (1) for the maximum sustainable κ_m for each carrier. The market's sustainable κ_m is the minimum across all legacy carriers.

While traditional methods estimate returns on capital from accounting data (Carlton and Perloff 2015), we instead infer possible values of r_{jt} using model restrictions.

Our econometric approach tests different values of κ_m using cost and demand specifications. For each κ_m, we calculate both the threshold discount factor $\underline{\delta}_{jmt}(\kappa_m)$ from equation (2) and the

[3]The stage game varies across time periods due to seasonal fluctuations and random shocks in both demand and costs. We primarily assume carriers treat current conditions as permanent when making choices (following Eizenberg et al. (2023).

[4]By specifying the problem in this way we ignore the role of multimarket contact. In principle that role could be introduced by aggregating the ICCs over markets where the same Legacy airlines overlap in the spirit of Eizenberg et al. (2023). Instead, we analyze the incentives in each market separately. Nonetheless, we would be able to explore whether coordination is stronger when multimarket contact is more important.

[5]See Eizenberg et al. (2023) for a proof of the existence and uniqueness of this κ value.

actual discount factor δ_{jt} from equation (3). We reject any κ_m where the implied discount factor falls below the implied threshold, leaving us with a set of non-rejected conduct values. We now specify the additional structure needed for this analysis.

2.2 Carriers' Technology and Costs

We now place additional structure on the variable profits $v(\cdot)$ that appear in the Incentive Compatibility Constraints (1). This is accomplished by modeling the airlines' technology, as well as the demand system they face. For the demand side, we rely on a Nested Logit specification that has been widely used in studies of airline markets (e.g., Ciliberto et al. (2021)). We leave a detailed description of this demand model to section 4.

Airlines require fixed proportions of capital (planes and infrastructure), labor (pilots and crew), and jet fuel for operations.[6] While a frictionless model would imply constant passenger volumes, airlines face both demand uncertainty and discrete capacity constraints in practice. Our model incorporates these key features.

Airlines use "revenue management" to optimize prices and capacity under demand uncertainty.[7] We capture this through the *load factor*, ϕ_{jmt}, the ratio of *actual* passengers to available seats.

We therefore determine that carrier j operating in market m in quarter t produces an output Q_{jmt} — the number of passengers flown — as follows:

$$Q_{jmt} = \phi_{jmt} \cdot \min\left\{a_j L_{jmt}, b_j F_{jmt}, K_{jmt}\right\}$$

where L is the number of stewards, K is capital and F is jet fuel in gallons, and a and b are carrier-specific parameters determining the proportions of labor and fuel necessary to operate the capital.

Capacity adjustments are costly for airlines. While they can modify seat configurations or change aircraft size, such changes incur significant costs. Moreover, staffing requirementslike the need for two pilots and a nonlinear relationship between cabin crew and seat countcreate additional rigidities.

The discrete nature of aircraft capacity leads to three cost categories in airline services (Panzar, 1979; Ciliberto et al., 2021): *airport* costs (sunk costs of gate service), *flight* costs (fuel and crew expenses), and passenger costs. Both airport and flight costs are fixed for a scheduled flight

[6] Bet (2021) models airline production with fixed fuel proportions but allows capital-labor substitution.

[7] A growing literature investigates the economic implications of these practices, e.g., Betancourt et al. (2022).

with a specific aircraft.[8] As these fixed costs don't affect pricing decisions, our marginal cost specification excludes fuel costs and salaries.

Passenger costs involve more than just the minimal accounting costs (like in-flight services) of an additional traveler. The true economic cost of a passenger is the opportunity cost, the highest alternative price at which that seat could be sold.[9]

Our approach, however, does not require us to obtain the passenger cost from accounting data, nor a precise modeling of the alternative cost. In light of the above discussion, we may write the marginal cost function as follows:

$$c_{jmt} = \phi_{jmt}(r_{jmt} + \gamma_{k(jm)t})p_{k(jm),t}, \tag{4}$$

which illustrates that the *economic opportunity cost* of transporting a passegner is a function of the load factor and of the *economic* cost of capital, which is given by $(r_{jmt} + \gamma_{ft}) \cdot p_{kt}$, where r_{jmt} is firm j's rate of return on capital, and $\gamma_{k(jm),t}$ and $p_{k(jm),t}$ are, respectively, the depreciation rate and the price of the physical capital (the aircraft) used in market m by airline j at time t.

2.3 An identified set for κ, r

We now show how to derive identified sets for the degree of collusion, and for the rate of return on capital.

The key observation is that, given a generic value for the conduct level κ_m, the model assumptions pin down the marginal cost c_{jmt} via two different channels. First, this marginal cost is obtained from the cost minimization assumption by equation (4). Second, the vector of market-m marginal costs is also implied by the static first-order conditions associated with Definition 1:

$$c_m(\kappa_m; \theta) = p_m - \left(\Omega^{\kappa_m} \odot \mathcal{S}(p; \theta)\right)^{-1} s(p; \theta), \tag{5}$$

where p_m is a vector containing the prices of all goods (flight options) in market m. To understand equation (5) note that first that it delivers the solution to the first-order conditions associated with each legacy carrier maximizing the sum of its own variable profits, and the variable profits of the other legacy carriers weighted by κ_m, while each low-cost carrier maximizes its own variable

[8] While airlines can modify seat configurations or change aircraft types, staffing requirements remain relatively rigid, requiring two pilots and similar crew sizes.

[9] This opportunity cost reflects potential demand from: future bookings on the same flight, alternative flights in the same market, and connecting flights across markets. Modeling these complex substitution patterns remains an open challenge in the literature.

profits only.

Since we only allow for single-product firms in each market, the matrix Ω^m has entries of 1 on its diagonal, and entries of κ_m at some of its off-diagonal elements: $\Omega^m_{jk} = \kappa_m$ if goods j, k are produced by *different legacy carriers*. The remaining off-diagonal elements obtain the value of zero. The remaining terms in (5) are defined as follows. The vector s contains market shares, determined as a function of market prices and the demand parameters θ. The matrix $\mathcal{S}$ contains market share derivatives such that $S_{jk} = -\partial s_k(p, \theta)/\partial p_j$. The latter are identified in a standard fashion given the nest logit demand system and data on prices, passengers flown and various cost and demand shifters (again, we delay discussion of the demand specification and its estimation to section 4).

Denote by $c_{jm}(\kappa_m; \theta)$ the component of the marginal cost vector in (5) pertaining to the marginal cost of Legacy carrier j. Since this is the true economic cost, it must be equal to the firm's market-m's marginal cost given in (4). By equating these costs we obtain a level of the rate of return on capital that is implied by the hypothesized level of coordination κ_m:

$$r_{jmt}(\kappa_m; \theta) = \frac{c_{jmt}(\kappa_m; \theta)}{\phi_{jmt} p_{k(jmt)}} - \gamma_{jmt} \tag{6}$$

Equation (6) motivates a thought experiment where one performs tests of specific coordination degrees. Each generic value of κ_m implies a rate of return on capital vie (6) which, via (3), implies a discount factor $\delta_j(\kappa_m)$. However, for κ_m to be sustainable, we must have $\delta_j(\kappa_m) \geq \underline{\delta}_{jm}(\kappa_m)$ per (2). It follows that certain values of κ_m may be rejected in the sense that they are not sustainable in equilibrium: they imply a discount factor that lies below the threshold discount factor necessary to sustain them. The model therefore restricts the set of coordination levels that are consistent with the model and the data.

To translate this intuition into a set identification result we must first state the set of observable quantities in market m, $\mathcal{X}_m$:

$$\mathcal{X}_m = \{\phi_{jm}, \gamma_{jm}, s_m, p_m, x_m, z_m\}, \tag{7}$$

where s_m is a vector of observed market shares, p_m is a vector of observed prices, x_m is a vector of observed demand shifters, and z_m is a vector of instruments that are correlated with prices but not with the demand unobservables ξ_m.

Given (s_m, p_m, x_m, z_m) the demand parameters θ are identified via standard arguments (Berry and Haile 2014). The market share derivatives $\mathcal{S}$ are then also identified. Along with the observed market shares and prices, and the known ownership matrix Ω^{κ_m}, the marginal cost vector $c_m(\kappa_m; \theta)$ and its component $c_{jm}(\kappa_m; \theta)$ are identified from (5) for any value of κ_m in the

$[0, 1]$ interval.

The rate of return $r_j(\kappa_m; \theta)$ is then identified from (6), and the discount factor $\delta_j(\kappa_m; \theta)$ is identified from (3). The threshold discount factor $\underline{\delta}_{fm}(\kappa_m, \theta)$ is identified from (2).[10]

The identified set for the degree of market-m coordination level κ_m is therefore given by:

$$\kappa_m^*(\mathcal{X}; \theta) = \{\kappa \in [0, 1] : \delta_{jm}(\kappa_m; \theta) \geq \underline{\delta}_{jm}(\kappa_m, \theta) \text{ for each } j \in L_m, \tag{8}$$

where L_m is market m's set of legacy carriers. The set $\kappa_m^*(\mathcal{X}; \theta)$ contains the coordination levels that are consistent with the data and the model. Specifically, these are the coordination degrees that are sustainable in some SPNE of the Supergame. Note that we also get an identified set of rate of return on capital for each legacy carrier j:

$$r_{jm}^*(\mathcal{X}; \theta) = \{r : r_j(\kappa_m; \hat{\theta}) \ \forall \kappa_m \in \kappa_m^*\}. \tag{9}$$

3 Data and Variable Definitions

Our study draws upon a diverse collection of datasets, including some that we believe to be original contributions. In the following section, we provide a detailed description of each dataset used in our analysis.

First, we use the Department of Transportation's Origin and Destination Survey (DB1B), a quarterly 10% sample of all tickets sold that offers data on ticket prices and passenger numbers for each carrier and route, as well as information for identifying connecting overlaps and controls. Crucially, we utilize the DB1B Coupon dataset, which provides detailed information on flight legs for each domestic itinerary. This dataset allows us to construct the sequence of routes, or legs, that carriers use to serve specific markets. For a detailed explanation of our sample selection process from the DB1B dataset, we direct readers to Appendix 7.1 which outlines our methodology in full.

Then, we utilize the Bureau of Transportation Statistics' T100 (Form 41) database to extract quarterly data on two key metrics: the number of allocated seats (capacity) by airline for each route and the number of passengers served on each route.

Next, we incorporate two additional datasets to identify the specific aircraft used on each route. We first utilize the "Airline On-Time Performance Data" to collect information on the

[10]This is based on the fact that given the demand parameters θ, prices can be calculated on the equilibrium path, in an optimal deviation from it, and in a competitive reversion, as counterfactuals. Given marginal costs, these quantities can be translated into the variable profits that appear in the Incentive Compatibility Constraint. See Eizenberg et al. (2023) for additional details.

tail number of each airframe used in a route. We then combine this with data from the "Air Carrier Financial: Schedule B-43 Inventory" to obtain detailed information about these aircraft. This combination allows us to determine crucial characteristics such as the manufacturer, year of manufacture, aircraft type, number of seats, and specific model.

Once we have the information on the aircraft and its age, we combine an original dataset that we have built using data from *VREF - Aircraft Value Reference & Aircraft Appraisals,* which we acquired for this project, to get the price of the aircraft, and the depreciation rate.

3.1 Airline Ticket Prices, Volume of Passengers, and Other Control Variables

Our analysis covers quarterly data from 2018Q1 to 2019Q4. We define a market-time observation (indexed by mt) as the combination of a year-quarter and a unidirectional airport-to-airport market.[11]

Our dataset covers markets between the top 50 Metropolitan Statistical Areas (ranked by 2016 population). It contains 114,140 carrier-market-time observations across 32,081 unique market-time combinations, where carriers are indexed by j and year-quarters by t.

Table 1: Summary Statistics

	Mean	Median	Std	Min	Max
Presence at Origin	0.85	0.92	0.37	0.01	1.52
Market Distance	1.256	1.088	0.685	0.067	2.719
Nonstop Percent	0.28	0.00	0.43	0.00	1.00
Market Share	0.001	0.00	0.00	0.00	0.04
Inside Share	0.28	0.18	0.29	0.00	1.00
Prices (hundreds)	2.08	2.11	0.60	0.51	4.73
Load Factor	0.83	0.84	0.06	0.29	1.00
Airframe Price per Seat (hundreds)	13.7	12.49	19.79	1.68	3436.05
Depreciation	-0.08	-0.08	0.01	-0.15	-0.04

Notes: Summary statistics from the DB1B, Form 41, and VREF datasets.

We calculate average fares, p_{jmt} at the carrier-market-time level.[12] As shown in Table 1, the

[11] We follow Borenstein (1989) and treat flights to different airports within the same metropolitan area as separate markets. For example, the ORD-LGA market in 2018Q2 includes both direct flights from Chicago O'Hare to New York La Guardia and connecting itineraries through cities like Cleveland. ORD-LGA and LGA-ORD constitute distinct markets.

[12] These fares are deflated using the second quarter of 2012 consumer price index to ensure comparability across different time periods.

average (uni-directional) fare in our sample is $208 with a standard deviation of $60.

On the demand side, we control for airlines' product differentiation using $\texttt{Presence at Origin}_{jmt}$, which measures the number of markets (in hundreds) that airline j serves from market m's origin airport at time t, including both nonstop and connecting flights (Berry, 1990, 1992; Ciliberto and Tamer, 2009). This network measure can be considered exogenous to prices in any given market-year, as establishing airline networks requires substantial long-term investments. On average, airlines serve 85 markets from each airport (Table 1). We also include $\texttt{Presence at Destination}_{jmt}$ as a cost shifter: while consumers are unlikely to value the number of connections available at their destination, a larger network at the destination airport presumably increases the opportunity cost of allocating seats to a given route.

We include $\texttt{NonStopPct}_{jmt}$ to capture consumers' preference for direct service, measuring the percentage of carrier j's passengers that are transported nonstop in market m at time t (Ciliberto and Williams, 2014). While the average share of nonstop flights is 28%, the median of 0% indicates that most markets are served only by connecting flights (Table 1).

We control for market distance using $\texttt{Market Distance}_{mt}$, measured in hundreds of miles. Additional control variables include carrier and year-quarter fixed effects.

3.2 Load Factors

The DB1B Coupon dataset provides detailed information on flight legs for domestic itineraries, allowing us to identify the sequence of routes carriers use to serve each market. A leg represents a *combination* of segments of a passenger's journey from origin to destination. For example, in the Charlottesville to Dallas Fort Worth (CHO-DFW) market, American Airlines might offer routing through either Charlotte (CHO-CLT-DFW) or Philadelphia (CHO-PHL-DFW). This leg-level data enables us to match aircraft and load factor information from other datasets.

We employ the Bureau of Transportation Statistics' T100 (Form 41) database to obtain quarterly data on two key metrics: allocated seat capacity by airline per route and actual passenger counts per route. From these metrics, we derive the load factor, ϕ_{jmt}, for each market, calculated as the ratio of passengers to available seats. The detailed methodology for aggregating this variable from route-level to market-level data is presented in Appendix 7.2.

The summary statistics in Table 1 reveal high capacity utilization in the airline industry, with an average load factor of 83 percent. The relatively small standard deviation of 6 percentage points suggests that load factors are fairly consistent across flights.

3.3 Airframe Prices and Depreciation Rates

We combine two additional datasets to identify the specific aircraft deployed on each route. First, we extract aircraft tail numbers from the "Airline On-Time Performance Data." We then link these identifiers to the "Air Carrier Financial: Schedule B-43 Inventory" to obtain detailed airframe characteristics. This matching process enables us to determine key aircraft attributes including manufacturer, production year, aircraft type, seat configuration, and specific model variant. Through this data integration, we can precisely identify the aircraft operating on each routefor example.

We augment our dataset with aircraft valuation data from VREF to determine both the price, $p_{k(jm)t}$, and depreciation rate, $\gamma_{k(jm)t}$, of each aircraft. The VREF dataset provides quarterly valuations stratified by manufacturing year. A detailed methodology for aggregating these variables from route-level to market-level observations is presented in Appendix 7.3.

As shown in Table 1, the average airframe price per seat is 1370$ with considerable variation across aircraft types as indicated by the standard deviation of 1979$. The analysis also reveals that airframes depreciate at a median annual rate of 8 percent, which captures the loss in value of aircraft over time.

4 Demand: specification and estimation results

We model consumer choice through a random utility framework over available flight options in each market. Our demand specification follows a nested logit structure (McFadden 1978, Berry 1994), which is widely established in airline industry demand analysis (Ciliberto, Murry, Tamer 2021). The model comprises two nests: one encompassing all domestic air service alternatives for the route, and another containing the outside optioneither foregoing travel on the route entirely or selecting alternative transportation modes such as rail or automobile.

The indirect utility of consumer i from choosing flight option j in market t is given by:

$$U_{ijmt} = \alpha p_{jmt} + \theta x_{jmt} + \xi_{jmt} + \bar{\epsilon}_{g(j)mt} + (1-\rho)\epsilon_{ijmt}, \tag{10}$$

where p_{jmt} is the price of flight option j in market m in quarter t, x_{jmt} is a vector of demand shifters, ξ_{jmt} is a carrier-specific unobservable, and $\bar{\epsilon}_{gjmt} + (1-\rho)\epsilon_{ijmt}$ is the familiar nested logit error term.[13] The parameter ρ captures the within-nest substitution, i.e., the degree of substitution between flight options. The parameter α captures price sensitivity, and θ is a vector

[13]Namely, the ϵ follows the Type I extreme value distribution and $\bar{\epsilon}$ has the unique distribution derived by Cardell (1997) that guarantees that the entire error term is also distributed Type I extreme value.

containing the remaining demand parameters.

Following Berry (1994), a market share equation can be inverted to obtain a linear estimating equation:

$$\ln(s_{jmt}) - \ln(s_{0mt}) = \alpha p_{jmt} + \theta x_{jmt} + \xi_{jmt} + \rho \ln(s_{j|g(j)mt}), \tag{11}$$

where $s_{j|g(j)mt}$ represents the share of product j within the nest.[14]

The term s_{0mt} denotes the market share of the outside option. To construct this measure, we first determine the market size using the geometric mean of origin and destination city populations, which is the standard approach in the literature.[15] We then calculate airline-specific market shares, s_{jmt}, for each carrier j in market m at time t by dividing passenger counts by the total market size. The outside share, s_{0mt}, captures the fraction of potential consumers who either choose not to travel or opt for alternative transportation modes, computed as the complement to the sum of all airline shares in the market.

Estimation is performed via IV-GMM using instruments that account for the simultaneity of demand and supply. The right-hand side of (11) contains two endogenous variables: the price and the within-nest share. We instrument for those endogenous variables using the fractions of passengers flown by the relevant carrier in both the origin and destination airports. Those account for airport presence, an established shifter of costs (Berry, 1992; Ciliberto and Tamer, 2009).

We implement two instrumental variable approaches. The first employs competitor characteristics at the airport level, following Ciliberto et al. (2021) and Aryal et al. (2022). The second uses functions of market competitors' characteristics, following Bresnahan (1987) and Berry et al. (1995). Both approaches yield comparable results, with our main specification adopting the first method. As discussed earlier, a key instrument in our analysis is the carrier's presence at destination airports.

Column 1 of Table 2 reports our baseline demand estimates. The model yields a price sensitivity parameter α of -1.37, carrying the expected negative sign. The nested logit correlation parameter ρ equals 0.25, falling within its theoretical bounds of $[0, 1)$. This magnitude indicates moderate substitutability between flight options, as higher values in this interval correspond to stronger substitution patterns among alternatives. We estimate both the correlation and price coefficients with high precision.

[14] As explained above, while carriers may serve market m with multiple offerings, we aggregate those such that each carrier j offers a single product in each market. These offerings define the observations j that enter the estimation equation (11).

[15] See, for example, Berry and Jia (2010).

Table 2: Demand estimates

	(IV)	(IV - Endogenous Nonstop)	(OLS)
Intercept	-6.814***	-6.890***	-7.811***
	(0.084)	(0.086)	(0.025)
Correlation (σ)	0.252***	0.260***	0.358***
	(0.006)	(0.006)	(0.003)
Prices (α)	-1.365***	-1.310***	-0.547***
	(0.046)	(0.047)	(0.009)
Percent Nonstop	1.973***	2.162***	1.943***
	(0.021)	(0.029)	(0.011)
Market Distance	0.594***	0.604***	0.236***
	(0.017)	(0.017)	(0.007)
Origin Presence	1.764***	1.628***	1.732***
	(0.017)	(0.022)	(0.018)
Carrier dummies			
AA	-0.604***	-0.544***	-0.911***
	(0.020)	(0.021)	(0.011)
Alaska	-0.054**	-0.151***	0.289***
	(0.021)	(0.023)	(0.019)
JetBlue	-0.079***	-0.153***	0.203***
	(0.021)	(0.022)	(0.017)
Delta	-0.361***	-0.315***	-0.718***
	(0.026)	(0.026)	(0.011)
Frontier	-1.663***	-1.657***	-0.573***
	(0.038)	(0.039)	(0.016)
Spirit	-1.436***	-1.466***	-0.386***
	(0.039)	(0.039)	(0.017)
United	-0.557***	-0.508***	-0.653***
	(0.017)	(0.017)	(0.008)
Year-quarter dummies			
2018Q1	0.058***	0.049***	0.021
	(0.017)	(0.017)	(0.017)
2018Q2	0.168***	0.164***	0.100***
	(0.017)	(0.017)	(0.016)
2018Q3	0.073***	0.067***	0.061***
	(0.016)	(0.016)	(0.016)
2018Q4	0.107***	0.101***	0.061***
	(0.016)	(0.016)	(0.016)
2019Q1	-0.024	-0.029*	-0.021
	(0.017)	(0.017)	(0.016)
2019Q2	0.182***	0.175***	0.111***
	(0.017)	(0.017)	(0.016)
2019Q3	0.090***	0.087***	0.068***
	(0.016)	(0.016)	(0.016)
Median Elasticity	-3.546	-3.433	-1.598
Number of Inelastic Products	42	79	15417
Observations	114140	114140	114140
R^2	0.681	0.686	0.726

Notes: Estimation of the nested logit demand equation (11). See text for variable definitions. *p<0.1; **p<0.05; ***p<0.01. Source: authors' estimates implied by the data and model assumptions.

The route's distance also has a positive and significant effect on utility as long distances favor flights over alternative modes of transportation.

The IV estimation yields a median own-price elasticity of demand of -3.546, consistent with findings in the literature. When we account for the endogeneity of nonstop flight percentages in Column 2, the elasticity remains nearly unchanged. However, failing to control for the endogeneity of both prices and $\ln(s_{j|g(j)mt})$ in Column 3 results in elasticity estimates roughly half as large. Notably, Column 3 reveals a much larger number of inelastic products compared to Columns 1 and 2.

A carrier's presence at the origin airport, $\texttt{Presence at Origin}_{jmt}$, significantly influences consumer choice. When airlines serve many routes from an origin airport, passengers can derive additional benefits through frequent flier programs. Our results confirm this relationship: Origin Presence shows a positive and significant effect on passenger utility, with an estimated coefficient of 1.764.

Route distance also positively affects utility, likely because air travel becomes increasingly advantageous over alternative transportation modes as distance grows. This relationship is captured by Market Distance, which shows that longer routes are associated with higher demand for air travel relative to not traveling (the outside good).

As expected, the variable $\texttt{NonStopPct}_{jmt}$ is associated with stronger demand for airline travel. The percentage of passengers flying nonstop on a specific carrier, measured by $\texttt{NonStopPct}_{jmt}$, is associated with higher demand for travel on that airline. This finding confirms that passengers value the convenience and shorter travel times of nonstop services, making carriers with higher proportions of nonstop flights more attractive.

Also estimated are the utility effects of carrier dummies (with Southwest being the omitted carrier) and dummy variables for each quarter between 2018Q1 and 2019Q3 (with 2019Q4 being the omitted quarter category). Both carrier and quarter effects are precisely estimated.

5 Supply: Exploring Coordination

Using our demand estimates, we implement the testing framework developed in Section 2 to examine whether different levels of market coordination could be sustained in equilibrium.

Our analysis centers on the conduct parameter $\kappa \in [0, 1]$, which measures the extent of coordination. For each strictly positive value of κ, we compare two key quantities: the discount factor implied by this conduct, and the threshold discount factor needed to sustain that conduct as a subgame perfect Nash equilibrium (SPNE).

We cannot reject a particular value of κ as the data generating process when the implied

discount factor exceeds the threshold. Testing a grid of κ values yields an estimated set of potentially sustainable coordination levels for each market, though we caution against interpreting these results as definitive evidence of coordination.[16]

Specifically, while $\overline{\kappa}_m$ represents the highest non-rejected coordination level in market m, this does not imply it was necessarily played, both because all smaller levels are also not rejected and because failure to reject merely indicates we cannot rule out $\overline{\kappa}_m$ as the true data generating process. Nevertheless, we view this statistic as a meaningful indicator of the potential for non-competitive conduct in market m.

Our initial findings reveal a systematic pattern in the estimated sets for the conduct parameter κ. In every market, these sets take the form of *intervals:* if $\overline{\kappa}_m$ represents the highest non-rejected coordination level in market m, then all values of κ below $\overline{\kappa}_m$ are also not rejected. This pattern is quite helpful in terms of describing our results. Each market m can now be simply characterized by its largest non-rejected coordination level $\overline{\kappa}_m$. We can therefore settle for analyzing the distribution of $\overline{\kappa}_m$ over markets.

Figure 1 shows how non-rejected coordination levels are distributed across markets. The results reveal a clear pattern: as κ increases, it is rejected in more markets. This pattern aligns with our interval-based estimated sets and, while intuitively sensible, was not guaranteed upfront.

Notably, perfect coordination ($\kappa = 1$) is only rejected in one-third of the markets. While these findings suggest significant potential for coordination across markets, we must interpret them cautiously for two reasons. First, failing to reject coordination does not prove its existence. Second, as Farrell and Baker (2021) argue, standard repeated games may overstate the ease of sustaining coordination through mechanisms like grim-trigger and deterministic punishment structures, rather than allowing for random coordination and punishment breakdowns.

To move beyond these limitations, we next examine *where* coordination is less readily rejected, allowing us to identify potential determinants of non-competitive behavior while netting out some model mis-specifications.

Potential facilitators. We next embark on a series of analyses that look for factors that correlate with non-rejection of various degrees of coordination. We specifically examine the role of familiar theoretical predictions: for example, that the presence of "mavericks," or non-symmetry in legacy carriers' market shares, may act as disruptors of coordination.

Examining correlates of coordination patterns provides valuable insights beyond our model's core structure, helping us understand features of airline markets that our framework does not explicitly incorporate. For instance, our analysis can reveal whether specific low-cost carriers

[16] We address cases where coordination would reduce profits for some legacy carriers by treating these carriers as Nash players and recalculating sustainable κ values for the remaining legacy carriers (when at least two remain).

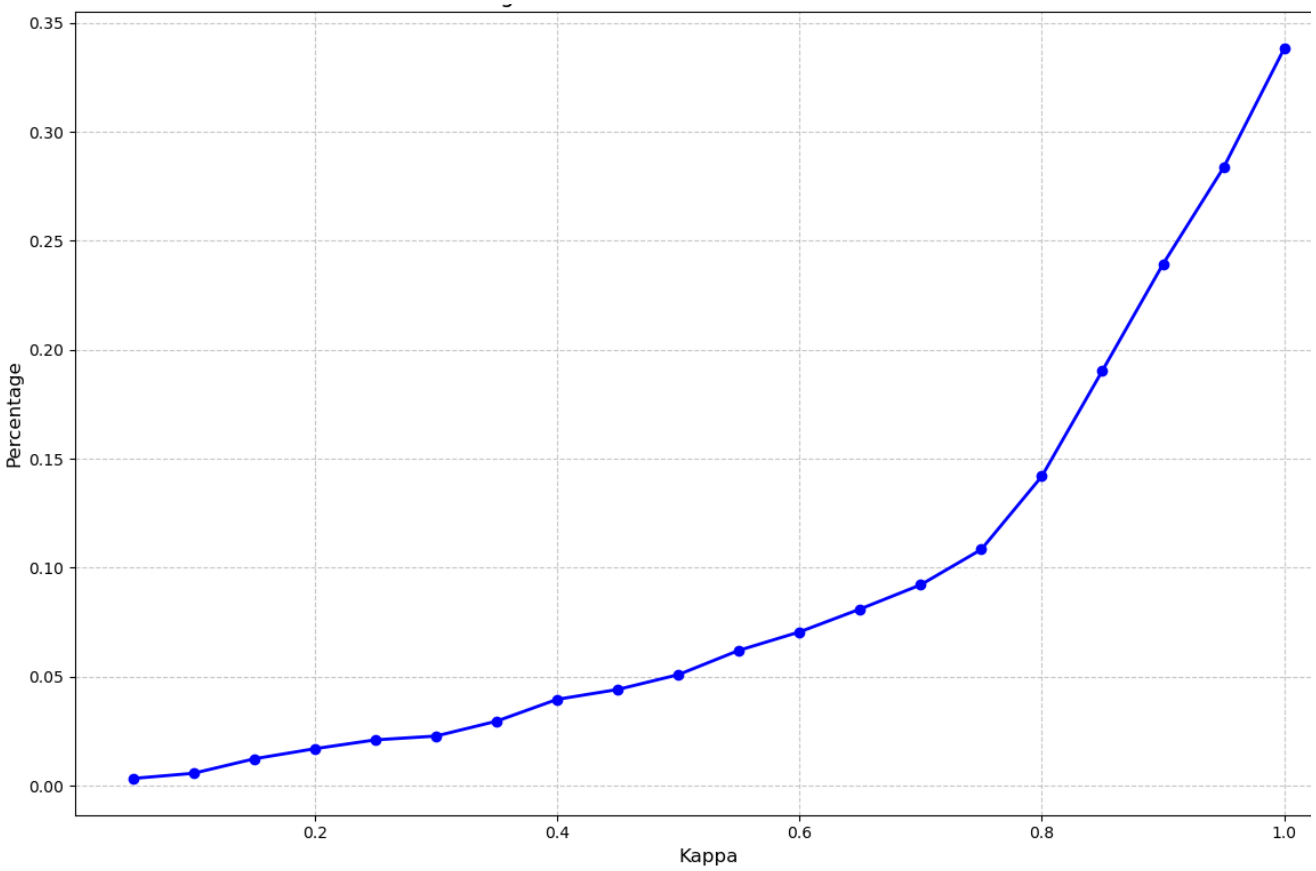


Figure 1: The fraction of markets where coordination of degree κ is rejected, by κ

disrupt coordination among legacy carriers, a finding that emerges from the data rather than being imposed by our model's structure.

We approach these correlations with appropriate caution: market structure features, including market shares and the presence of low-cost carriers, arise endogenously rather than randomly. Additionally, since our framework incorporates coordination theory, some theoretical predictions may be embedded in our model. This limits our ability to conduct fully independent tests of these predictions, but does not diminish the value of identifying which specific predictions are supported by our estimated coordination levels.

Our analysis proceeds in two stages. First, we use heatmaps to visualize correlation patterns non-parametrically, allowing us to detect potential non-linear relationships. We then complement this analysis with regressions to quantify these relationships more formally.

Figure 2 explores whether market share symmetry among legacy carriers facilitates coordination. The heatmap displays rejection rates for different coordination levels (κ) on the horizontal axis and legacy carriers' Mean Absolute Deviation in market shares on the vertical axis. Higher

positions on the vertical axis indicate greater market share symmetry among carriers.

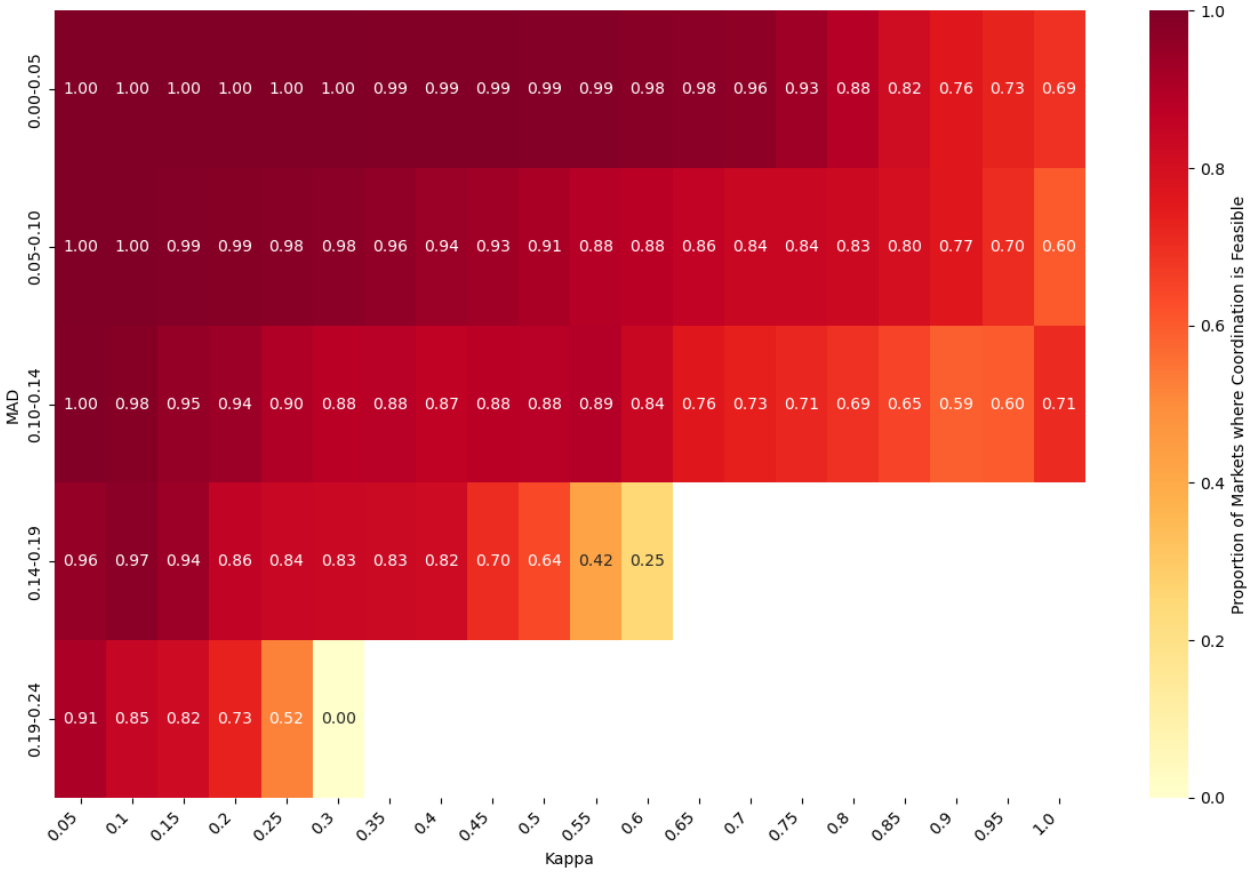


Figure 2: The fraction of markets where coordination of degree κ is not rejected, by κ and the Mean Absolute Deviation of legacy carriers' market shares

The pattern is striking: as we move down, we see that for almost all levels of κ, coordination is more frequently rejected when market share asymmetry is higher. This pattern aligns well with theoretical predictions regarding the role of symmetric players in fostering stable coordination.

Figure 3 examines how the presence of Low Cost Carriers (LCCs) relates to coordination patterns. For coordination levels above $\kappa = 0.2$, markets with LCCs show higher rejection rates of coordination. This effect is particularly strong for high coordination levelsfull coordination cannot be rejected in 83% of markets without LCCs, compared to only 59% of markets with LCCs. These findings suggest LCCs may act as potential disruptors of non-competitive equilibria.

To understand which LCCs drive this effect, we examine specific carrier types. Figure 4 focuses on "maverick" carriers (Alaska and JetBlue), but finds they have little to no effect on coordination rejection rates. Similarly, Figure 5 shows that Ultra Low Cost Carriers (Spirit, Frontier, Sun Country) have only modest effects.

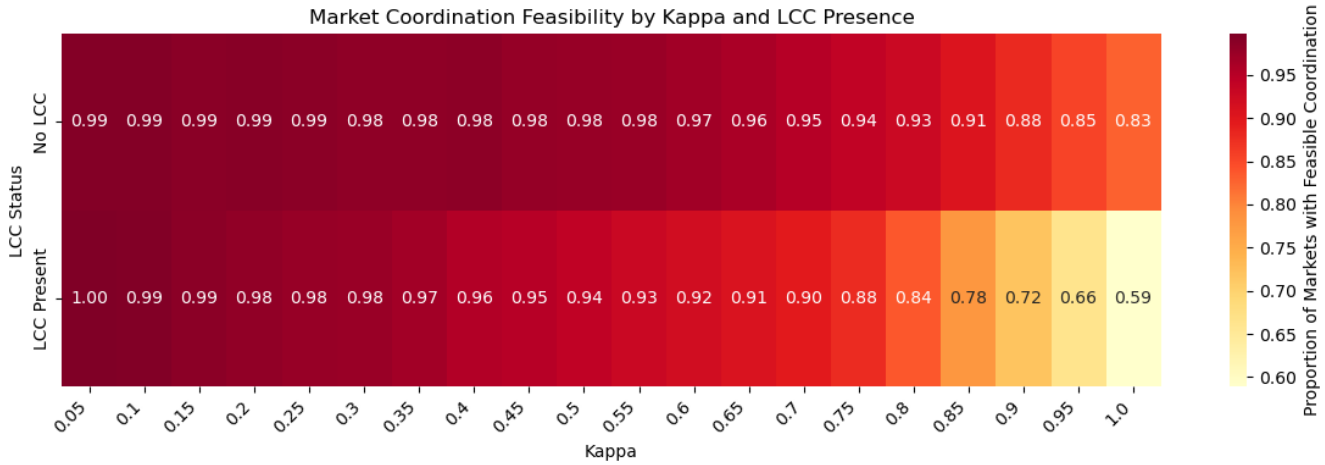


Figure 3: The fraction of markets where coordination of degree κ is not rejected, by κ and LCC presence

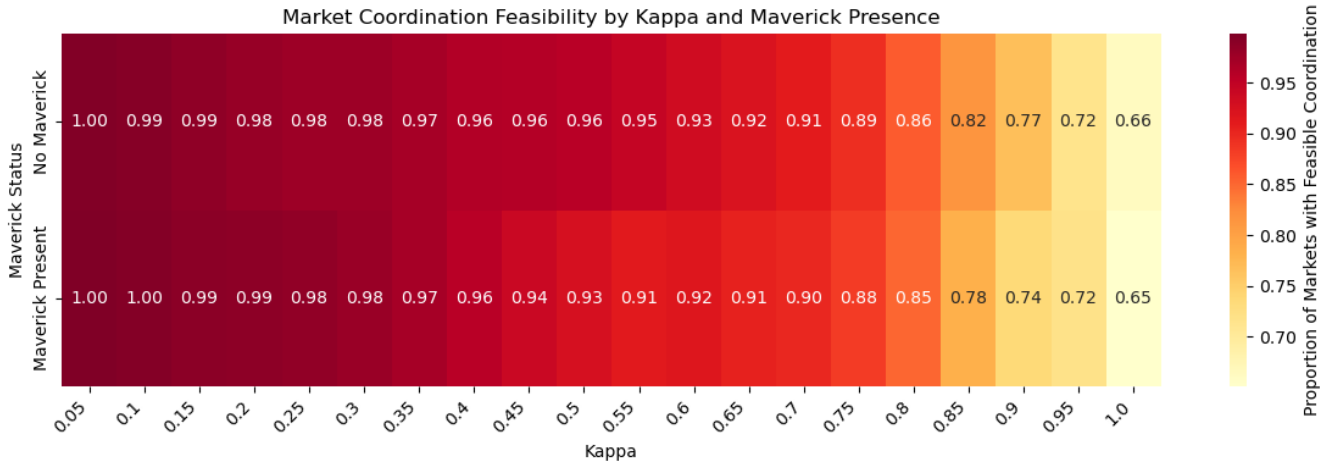


Figure 4: The fraction of markets where coordination of degree κ is not rejected, by κ and Maverick presence (JetBlue,Alaska)

The disrupting effect appears to be driven primarily by Southwest Airlines, as shown in Figure 6. Markets with Southwest present show substantially higher rejection rates of coordinationfull coordination cannot be rejected in 79% of markets without Southwest, compared to 57% of markets with Southwest.

Table 3 quantifies the relationships between LCC presence and coordination through Probit regressions. Our dependent variable equals 1 when full coordination ($\kappa = 1$) is not rejected in a

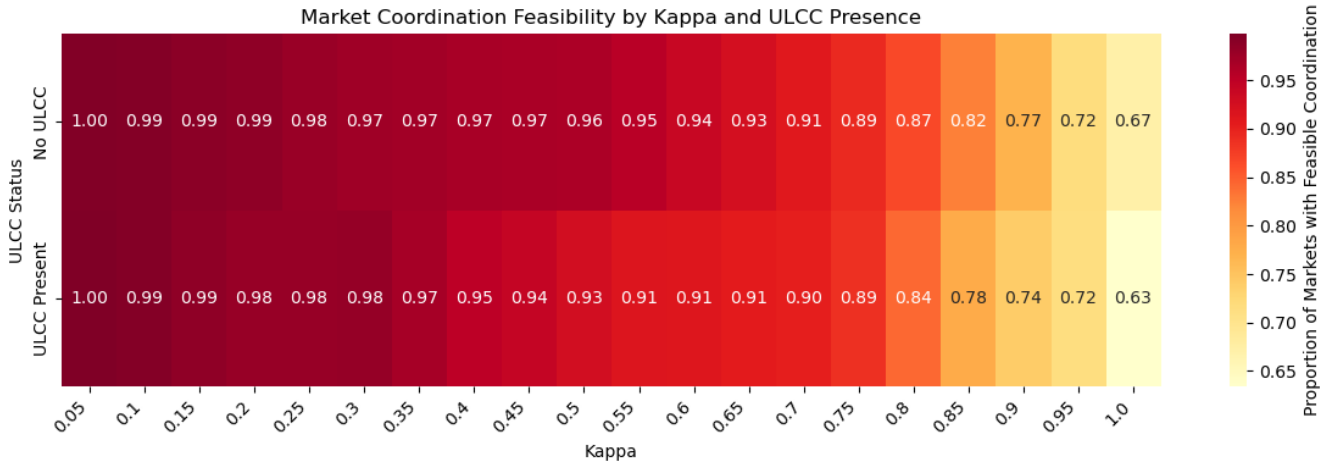


Figure 5: The fraction of markets where coordination of degree κ is not rejected, by κ and Ultra Low-Cost Carrier presence (Spirit, Frontier, Sun Country)

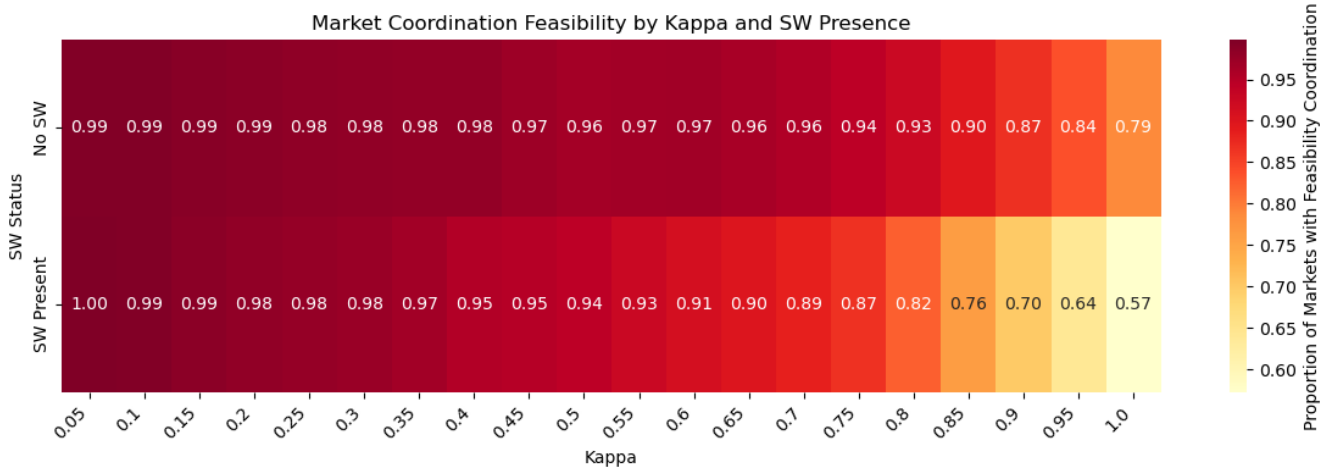


Figure 6: The fraction of markets where coordination of degree κ is not rejected, by κ and the presence of Southwest Airlines

market during 2018Q1-2019Q4.

The first specification confirms our earlier heatmap findings: LCC presence significantly reduces the likelihood of non-rejected full coordination. Specification (2) breaks down this effect by carrier type, revealing Southwest Airlines as the primary driver. Southwest's negative effect on non-rejected coordination is approximately 12 times larger than that of Ultra LCCs. Interestingly, "maverick" carriers (Alaska and JetBlue) show a small positive correlation, suggesting

Table 3: Probit analysis of non-rejection of full coordination

Maverick		0.061***
		(0.022)
LCC	-0.726***	
	(0.018)	
Southwest		-0.615***
		(0.016)
ULCC		-0.049***
		(0.018)
Intercept	0.952***	0.803***
	(0.026)	(0.024)
2018Q2	0.092***	0.080***
	(0.031)	(0.031)
2018Q3	0.121***	0.126***
	(0.031)	(0.031)
2018Q4	0.113***	0.111***
	(0.031)	(0.031)
2019Q1	-0.145***	-0.140***
	(0.030)	(0.030)
2019Q2	-0.116***	-0.112***
	(0.030)	(0.030)
2019Q3	-0.075**	-0.077**
	(0.030)	(0.030)
2019Q4	0.025	0.014
	(0.030)	(0.031)
Observations	29580	29580
Pseudo R^2	0.050	0.045

Notes: Probit regressions where the dependent variable is a binary variable taking the value 1 if, in the relevant market, full coordination (i.e., κ=1) has not been rejected, and zero otherwise. LCC is an indicator for the presence of any low cost carrier (i.e., non-legacy) in the market. ULCC is an indicator for the presence of an ultra-low cost carrier (Spirit, Frontier, Sun Country). Maverick is an indicator for the presence of JetBlue or Alaska. Southwest is an indicator for the presence of Southwest airlines. *p<0.1; **p<0.05; ***p<0.01. Source: authors' estimates implied by the data and model assumptions.

they do not impede coordination among legacy carriers.

Table 4 extends our analysis using OLS regressions where the dependent variable is the market's highest non-rejected coordination level ($\overline{\kappa}_m$). Specification (1) shows that LCC presence significantly reduces this maximum κ by 0.19.

When examining specific carrier types in specification (2), all LCCsUltra LCCs, mavericks, and Southwestshow similar negative effects, with coefficients ranging from -0.07 to -0.12. While Southwest maintains the strongest negative correlation, its distinctive effect is less pronounced than in our earlier Probit analysis of full coordination. This suggests that Southwest's disruptive role is particularly strong for high coordination levels.

Taking our findings together, Southwest Airlines emerges as the strongest correlate with reduced coordination among legacy carriers, especially at high coordination levels. However, when examining the full range of coordination values, other LCCs also show meaningful negative correlations with coordination levels.

6 Concluding remarks

We develop and estimate a structural model of coordination among legacy carriers in the US airline industry during 2018-2019. Our primary contribution is deriving market-specific identified sets for the degree of coordination by exploiting a key insight: the unobserved rate of return on capital determines both carriers' markups and their patience to sustain these markups in equilibrium. To implement this approach, we construct a novel dataset linking aircraft replacement costs to market-specific airframe usage, enabling us to connect markups with coordination sustainability.

The estimated sets take an interval form, with the largest non-rejected degree of coordination serving as our key measure. Our analysis reveals a substantial potential for coordinationwe cannot reject full coordination in about two-thirds of markets. While this high level of non-rejected coordination may reflect limitations in our repeated game framework, comparing across markets allows us to net out some of these model-specific effects.

These comparisons yield patterns that are consistent with theoretical predictions: coordination is less likely to be sustained when low-cost carriers are present and when legacy carriers have asymmetric market shares. Southwest Airlines emerges as particularly importantits presence shows the strongest negative correlation with coordination, especially at its high levels.

Importantly, we must interpret these correlations carefully. The presence of low-cost carriers is endogenousstrong coordination among legacy carriers might attract entry. Moreover, our conclusions about non-rejected coordination levels come from a model that already accounts

Table 4: Determinants of largest non-rejected coordination level

Intercept	0.819***	0.797***
	(0.007)	(0.007)
Maverick		-0.079***
		(0.005)
LCC	-0.193***	
	(0.005)	
Southwest		-0.116***
		(0.004)
ULCC		-0.107***
		(0.004)
2018Q2	-0.017**	-0.010
	(0.008)	(0.008)
2018Q3	-0.016*	-0.007
	(0.008)	(0.008)
2018Q4	-0.024***	-0.017**
	(0.008)	(0.008)
2019Q1	-0.006	-0.004
	(0.008)	(0.008)
2019Q2	-0.023***	-0.014*
	(0.008)	(0.008)
2019Q3	-0.020**	-0.010
	(0.008)	(0.008)
2019Q4	-0.020**	-0.010
	(0.008)	(0.008)
Observations	25085	25085
Adjusted R^2	0.059	0.080

Notes: Linear regressions estimated via OLS where the dependent variable is the largest value of κ that has not been rejected in the relevant market. LCC is an indicator for the presence of any low cost carrier (i.e., non-legacy) in the market. ULCC is an indicator for the presence of an ultra-low cost carrier (Spirit, Frontier, Sun Country). Maverick is an indicator for the presence of JetBlue or Alaska. Southwest is an indicator for the presence of Southwest Airlines. *p<0.1; **p<0.05; ***p<0.01. Source: authors' estimates implied by the data and model assumptions.

for low-cost carriers' presence. Yet, while our model's assumptions might drive some general patterns, they are unlikely to explain specific findings like Southwest's distinctive role.

7 Appendix: Sample Construction from the DB1B

7.1 DB1B Dataset Preparation

Our main data source is the domestic Origin and Destination Survey (DB1B), a 10% sample of airline tickets from all reporting carriers. This dataset, available from the Department of Transportation's National Transportation Library, has three versions:

- DB1B Coupon: The basis of our analysis, providing coupon-specific information for each domestic itinerary.
- DB1B Market: Contains directional market characteristics for each Origin-Destination domestic itinerary.
- DB1B Ticket: Includes information on reporting carrier, itinerary fare, number of passengers, originating airport, roundtrip indicator, and miles flown.

We clean the data by removing:

- Tickets with more than 6 coupons overall, or more than 3 coupons in either direction for round-trips.
- Tickets sold by foreign carriers flying between two US points.
- Tickets that are part of international travel.
- Tickets involving non-contiguous domestic travel (Hawaii, Alaska, and Territories).
- Tickets with questionable fare credibility or bulk fare indicators.
- Tickets that are neither one-way nor round-trip.
- Interline tickets (travel on more than one airline in a directional trip).

Following Borenstein (1989) and Ciliberto and Williams (2014), we treat roundtrip tickets as two directional trips, with each direction's fare being half of the roundtrip fare. A passenger flying on a roundtrip ticket is counted as a single passenger in both directions. We exclude tickets with fares below \$25, those in the top and bottom 1% of the year-quarter fare distribution, and airlines serving fewer than 90 passengers per quarter.

7.2 Market Load Factor

There is no data available on the load factors at the market level. Data are only available at the (nonstop) route level. Constructing the load factor at the market level from the route level presents several challenges due to the complex relationships between markets, legs, and routes, as well as the nature of the available datasets. A single market can often be served via multiple legs, and each leg can include various routes. Furthermore, any given route within a leg typically serves multiple markets.

To address these intricacies, we employ a weighted average approach. This method allows us to calculate an average load factor for each market that is internally consistent while ensuring that the total load factor aligns with the actual observed values. By carefully considering the contribution of each route to the overall market, we can allocate an appropriate portion of the load factor to each market.

This approach maintains the integrity of the data across different levels of aggregation, from individual routes to entire markets, providing what we believe is an accurate representation of airline capacity utilization within our analytical framework.

To illustrate our practical approach, let's revisit the example from the main text: the Charlottesville to Dallas Fort Worth (CHO-DFW) market. American Airlines offers two routing options: CHO-CLT-DFW via Charlotte or CHO-PHL-DFW via Philadelphia. We use two primary data sources:

The Bureau of Transportation Statistics' T100 (Form 41) dataset provides load factors for the individual routes: CHOCLT, CLTDFW, CHOPHL, and PHLDFW.

The DB1B Coupon dataset, a 10% sample, provides information on:

- The number of passengers transported by American on each route (e.g., CHOCLT) who were flying from CHO to DFW.
- The total number of passengers on each route.

Ideally, the total passenger count from DB1B should match the T100 dataset, but discrepancies exist due to DB1B being a sample. Our process involves:

- Calculating a "weight" for each route based on the proportion of CHO-DFW passengers to total passengers on that route.
- Multiplying this weight by the route's load factor from the T100 dataset.
- Repeating this process for all routes involved in the market.

- Computing the average load factor for each leg (e.g., CHO-CLT and CLT-DFW).
- Finally, calculating the average load factor for the entire market across all legs.

This method allows us to estimate market-level load factors while accounting for the complexities of multi-leg routes and varying passenger distributions.

7.3 Airframe Prices and Depreciation Rates

The next step in our process involves computing the average price and depreciation rates of the planes used in each market. This calculation follows a similar algorithm to the one used for load factors, but with two key differences:

- We account for the fact that airlines use multiple airframes each day, as reported in the "Airline On-Time Performance Data" dataset. To address this, we first calculate an average value of prices and depreciation rates for each route on a quarterly basis.
- Instead of using the T-100 dataset, we source our information from VREF. This dataset provides the necessary details on aircraft prices and depreciation rates.

Our method maintains the weighted average approach used in the load factor calculations. We consider the proportion of passengers on each route within a market, using this as a weight to allocate the appropriate share of aircraft prices and depreciation rates to each market. This approach allows us to estimate market-level aircraft characteristics while accounting for the complexity of airline fleet management and route structures.

7.3.1 Airframe Price Scraping and Imputation

To construct the airframe price dataset, we first downloaded PDF reports on each aircraft model available in both the Frames Data and in VREF. We then scraped the data from these reports to obtain the price of each aircraft model in each quarter. This process involved extracting the price information from charts in the PDFs and converting it into a structured dataset.

After this extraction process, we used Bayesian multiple imputation to construct prices for the missing aircraft (approximately 40 percent of the sample).

References

Abreu, D. (1986). Extremal equilibria of oligopolistic supergames. *Journal of Economic Theory*, 39(1):191–225.

Aryal, G., Ciliberto, F., and Leyden, B. T. (2022). Coordinated capacity reductions and public communication in the airline industry. *The Review of Economic Studies*, 89(6):3055–3084.

Berry, S. and Jia, P. (2010). Tracing the woes: An empirical analysis of the airline industry. *American Economic Journal: Microeconomics*, 2(3):1–43.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890.

Berry, S. T. (1990). Airport presence as product differentiation. *The American Economic Review*, 80(2):394–399.

Berry, S. T. (1992). Estimation of a model of entry in the airline industry. *Econometrica: Journal of the Econometric Society*, pages 889–917.

Betancourt, J. M., Hortaçsu, A., Oery, A., and Williams, K. R. (2022). Dynamic price competition: Theory and evidence from airline markets. Technical report, National Bureau of Economic Research.

Borenstein, S. (1989). Hubs and high fares: dominance and market power in the us airline industry. *The RAND Journal of Economics*, pages 344–365.

Bresnahan, T. F. (1987). Competition and collusion in the american automobile industry: The 1955 price war. *The Journal of Industrial Economics*, pages 457–482.

Bresnahan, T. F. (1989). Empirical studies of industries with market power. *Handbook of industrial organization*, 2:1011–1057.

Ciliberto, F., Murry, C., and Tamer, E. (2021). Market structure and competition in airline markets. *Journal of Political Economy*, 129(11):2995–3038.

Ciliberto, F. and Tamer, E. (2009). Market structure and multiple equilibria in airline markets. *Econometrica*, 77(6):1791–1828.

Ciliberto, F., Watkins, E., and Williams, J. W. (2019). Collusive pricing patterns in the us airline industry. *International Journal of Industrial Organization*, 62:136–157.

Ciliberto, F. and Williams, J. W. (2014). Does multimarket contact facilitate tacit collusion? inference on conduct parameters in the airline industry. *The RAND Journal of Economics*, 45(4):764–791.

Eizenberg, A., Shilian, D., and Blanga, D. D. (2023). Estimating the strategic effect of multimarket contact. Technical report, mimeo.

Farrell, J. and Baker, J. B. (2021). Natural oligopoly responses, repeated games, and coordinated effects in merger analysis: A perspective and research agenda. *Review of Industrial Organization*, 58(1):103–141.

Friedman, J. W. (1971). A non-cooperative equilibrium for supergames. *The Review of Economic Studies*, 38(1):1–12.

Goto, U. and Iizuka, T. (2016). Cartel sustainability in retail markets: Evidence from a health service sector. *International journal of industrial organization*, 49:36–58.

Green, E. J. and Porter, R. H. (1984). Noncooperative collusion under imperfect price information. *Econometrica: Journal of the Econometric Society*, pages 87–100.

Igami, M. and Sugaya, T. (2022). Measuring the incentive to collude: The vitamin cartels, 1990–99. *The Review of Economic Studies*, 89(3):1460–1494.

Kim, S. J. and Park, Y. (2023). Examining the coordinated effects of the aa/usair merger.

Miller, N. H., Sheu, G., and Weinberg, M. C. (2021). Oligopolistic price leadership and mergers: The united states beer industry. *American Economic Review*, 111(10):3123–59.

Panzar, J. C. (1979). Equilibrium and welfare in unregulated airline markets. *The American Economic Review*, 69(2):92–95.

Porter, R. H. (1983). A study of cartel stability: the joint executive committee, 1880-1886. *The Bell Journal of Economics*, pages 301–314.

Schmalensee, R. (1989). Inter-industry studies of structure and performance. *Handbook of Industrial Organization/in R. Schmalensee and R. Willig, Eds, Handbook of Industrial Organization, North Holland, New York*.

Turner, D. (2021). Coordinated effects in the american airlines-us airways merger. *Available at SSRN 3917112*.