A Game-Theoretic Modeling Approach to Air Traffic Forecasting

Vikrant Vaze, Reed Harder Thayer School of Engineering, Dartmouth College Hanover, NH 03755 vikrant.s.vaze@dartmouth.edu

*Abstract***—Federal, state and local aviation planners rely on air traffic forecasts for workforce staff planning (particularly for air traffic controllers), evaluation of current and future technological improvements at airports, planning of airport capacity expansion, and evaluation of federal funding requests for airport infrastructure improvements. While most existing forecasting models are econometric or statistical in nature, incorporating a more behavioral understanding of airline competition into the forecasts and planning process represents a significant opportunity to improve their efficacy. To this end, we develop a two-stage game-theoretic model for predicting airline capacity allocation decisions under competition. We first demonstrate desirable theoretical properties and computational tractability of our model, and then exploit them to develop solution algorithms with very fast convergence properties to enable rapid generation of forecasts, requiring only a few seconds of run-time. We then solve our model to equilibrium of the two-stage game using a real-world dataset based on an 11-airport, four-airline network from the western United States. The out-of-sample validations of traffic predictions at the airport and OD (origin-destination) level indicate a good fit to real-world traffic data for various lookahead times and at various levels of granularity. We thus demonstrate that a behaviorally consistent two-stage game model of airline competition provides a good fit to observed airline operations at various levels of aggregation, in turn highlighting its potential for accurate air traffic forecasting.**

Keywords- Air Traffic Forecasts; Game Theory; Airline Competition; Nash Equilibrium; Frequency and Fare Competition.

I. INTRODUCTION

Air traffic growth in an airport, city or region benefits the neighboring communities by enhancing access, by improving passenger convenience, and by stimulating economic activity. On the other hand, unforeseen traffic growth accompanied by insufficient planning to handle this growth can lead to worsening of undesirable outcomes, including congestion, delays, noise and pollution. Therefore, anticipating and proactively responding to changes in air traffic volumes is an important priority for aviation planners at the airport, city, state and federal levels. To this end, the Federal Aviation Administration's (FAA) official forecast of the U.S. aviation activity, called TAF-M (Modernized Terminal Area Forecast), combines statistical and econometric models to forecast flows of passengers and aircraft at various levels including OD market level, route level, segment level and airport level [1]. The reader is referred to [2] for a brief review of the major statistical approaches to air traffic prediction, as used within the FAA and outside. The TAF-M tool is used to understand the impacts of NEXTGEN on airports, passenger routings, and aircraft networks. NEXTGEN (Next Generation Air Transportation System) here refers to the FAA's vision and associated technologies and procedures for transforming United States' National Airspace System (NAS) in general, and the air traffic control in particular. TAF-M is used in a variety of ways to help in understanding the effects of policies, procedures and environmental regulations underlying NEXTGEN. For example, within the FAA, TAF-M forecasts are used by the Office of Airports (ARP) for evaluating airport investments, by the Air Traffic Organization (ATO) for NASwide simulations using airport-level operations, and by the Office of Energy and Environment (AEE), for evaluating noise and pollution at the airport and en route. Currently TAF-M is primarily an unconstrained demand forecasting tool that lacks prediction of airline behavioral response which is needed to predict changes in their flight frequencies. Thus, there exists a need to develop behavioral models of airline decision-making to augment and enhance current forecasting tools. This is the aim of our research.

Airline decisions are made in a competitive environment, and can be divided into capacity and fare decisions. Capacity decisions, including decisions about seats-per-flight and service frequency, affect both the operating costs and revenues of airlines. These decisions have significant implications for the performance of the air transportation system as a whole: overand under-allocation of airline capacity has been shown to result in billions of dollars in additional costs to airlines and passengers, wastage of system resources, passenger inconvenience and environmental damages. Airline frequency competition in particular has been shown to be a major driver of increased airport congestion. Prediction of the frequency component of capacity allocation decisions is of particular interest to us for a variety of reasons. First, unlike seats-perflight decisions, these decisions significantly affect schedule attractiveness to passengers, because frequency allows passengers greater scheduling flexibility. Additionally, these decisions show far greater variability across time and across the network than do the seating capacity decisions [3]. Therefore, in this paper, we will focus on predicting only the frequency component of capacity allocation decisions, while leaving predictions of seats-per-flight as a potential direction for future research.

Frequency and fare decisions of an airline are dependent on each other, and hence neither can be modeled in isolation. A higher frequency, for example, typically increases passenger attractiveness, thus enabling the airline to sell tickets at a higher price. Additionally, decisions of different airlines are interdependent as well: a higher frequency or a lower fare than

competing airlines can typically attract more passengers to an airline, while reducing the demand for the competitors. These interdependencies are well-captured by game-theoretic models.

Decisions about frequency and fare are typically made sequentially, on different timelines. Frequency decisions are often made weeks or months in advance of the flights in question, with only an approximate understanding of future fare decisions. On the other hand, fare decisions can be made days or even minutes in advance, with exact knowledge of frequency decisions. Moreover, these two kinds of decisions are made by very different departments within an airline that typically do not jointly optimize their decisions. Therefore, two-stage game-theoretic models, rather than single-stage game-theoretic models, are much more behaviorally adequate for describing these two kinds of decisions by the airlines.

In summary, the frequency and fare decisions are interdependent and are important factors affecting the performance of the air transportation system as a whole. Therefore, it is important to develop tractable models and solution concepts that accurately describe their dynamics. In this paper, we present a two-stage game-theoretic model of airline competition, demonstrate its tractability across a range of assumptions and parameter values, and validate its predictions against observed airline behavior. This two-stage game approach accounts for both the interdependence of competing airlines' behaviors and the sequential nature of frequency and fare decisions by the airlines as observed in practice. Each airline is assumed to pick a frequency value in each segment during the first stage of the game, in order to maximize its own profit. In the second stage, each airline decides the fare to be charged in each market, again while maximizing its own profit.

In this paper, we first present desirable theoretical properties, namely concavity and submodularity, for a simplified version of our game model. These properties are vital to ensure the availability of fast solution approaches for our model, especially since large-scale game-theoretic models can be very difficult to solve without such properties. We then extend our model by relaxing several of our assumptions, and demonstrate numerically that for a wide range of realistic values of model parameters, concavity and submodularity properties hold in an approximate way. These analytical results enable us to employ a tractable solution algorithm for finding an equilibrium of this game, and ensure that it converges rapidly. We use this algorithm to generate forecasts for a realworld case study network consisting of four major airlines making frequency decisions across a network of 11 airports in the western United States. The frequency forecasts from our equilibrium solution are then validated against the observed frequencies of these airlines over the same period. We then examine frequency predictions at various levels of aggregation,
including individual airline-segment pairs, individual including individual airline-segment pairs, segments, individual airports, individual airlines, airlinesegment types, and the overall network for look-ahead horizons at various time scales, and find that our predictions approximate airline behavior with good accuracy. Decisionmakers within the air transportation system may thus find refinements of our model useful for forecasting and scenario

analysis at various scenario granularities of practical importance.

II. AIRLINE COMPETITION BASICS

Profitability maximization is typically considered to be the primary objective behind an airline's decisions, and an airline's passenger share in a market is an important determinant of its profitability. Airlines strive to gain market share by offering itineraries that are most attractive to passengers. Itinerary attractiveness is a function of attributes such as fare, departure time, itinerary elapsed time, number of stops, connection time, as well as some other factors such as baggage fees, frequent flyer programs, on-board amenities, etc. Fare and schedule convenience are widely acknowledged to be the two most important attributes that affect passenger itinerary choice, and flight frequency (or the number of flights per day on a nonstop segment) is considered to be the most important dimension of schedule convenience. With more frequency, more passengers are likely to find an itinerary whose departure and arrival times match the passenger's travel preferences. However, the likelihood of a passenger choosing a particular itinerary is dependent on the attributes of that itinerary as well as the attributes of other itineraries, including itineraries of the other airlines, in that same market. Therefore, many existing studies have taken a game-theoretic approach to modeling airline competition on decisions related to schedule convenience and fares.

Most existing studies on airline competition have focused on formulating and solving game-theoretic models using the concept of Nash equilibrium or one of its extensions. Some studies have solved realistic-sized case study instances using successive optimization based algorithms, while others have proved desirable analytical properties of these games using simplified (*stylized*) models. However, most existing studies do not attempt to use these game models for predictions and forecasting. Others which do use them for forecasting show mixed results in terms of prediction accuracy. In one of the first studies in airline frequency competition, Hansen solved a frequency competition game using a successive optimizations algorithm for a network including 52 U.S. airports and 28 airlines, but model predictions showed some significant deviations from the empirical data [4]. Adler modeled airline competition on network construction in the first stage, and on frequency, seats, and fares in the second stage, but did not provide any results on solution tractability or empirical validation [5] [6]. Adler, Pels, and Nash solved a single-stage frequency, seats, and fares game, but did not validate the results empirically [7]. In a series of studies, Vaze and Barnhart studied single-stage frequency-only game [8] [9] [10], and found (in [8]) reasonable agreement between equilibrium predictions and observed frequencies. Multiple prior studies (e.g., [11], [12], [13]) formulated and solved a single-stage frequency-fare game, but did not provide any empirical validation of their results. Wei and Hansen solved a simplified single-stage frequency and seat allocation game through enumeration [14], while Brueckner analytically solved a simplified model of single-stage game of frequency, seats, and fares [15].

Very few past studies have focused on two-stage frequency-fare games, which is the focus of present research. First, Dobson and Lederer used heuristics for solving a twostage frequency-fare game [16]. Then Schipper, Rietveld, and Nijkamp analyzed the shift from monopoly to duopoly following airline deregulation by simulating a two-stage frequency-fare game [17]. Then, Brueckner and Flores-Fillol analytically compared the properties of the two-stage frequency-fare game with a single-stage frequency-fare game [18]. All three of these studies focused only on the simple case of one market with two airlines. None of these studies attempted to solve the game for real-world networks, nor did they empirically validate their models. Recently, Hansen and Liu noted the difficulty of solving two-stage games analytically, and instead just presented a small numerical example [13]. Several studies have stressed the need to develop two-stage frequency-fare game-theoretic models in order to account for the sequential nature of these decisions (e.g., [13], [16], [17], [18], [19]), but to our knowledge, no study has successfully bridged analytical, computational, and empirical approaches for such models. This is the goal of our present research project. In this particular paper, we will briefly describe our analytical and computational results, but focus specifically on the empirical validation and forecasting applications. For more detailed discussion of the analytical and computational components, the reader is referred to [3].

Two-stage game models, while behaviorally consistent, present several practical challenges. The major solution concept available for two-stage games, that of subgame perfect Nash Equilibrium, can be difficult to analyze mathematically, and intractable to compute in practice. In general, Nash equilibria may not exist for certain games, or multiple equilibria may exist in certain others. Even when equilibria do exist, they can be prohibitively expensive to compute even for simple models, let alone for extended manyplayer networks (e.g., [20]). Furthermore, game-theoretic models in this area can be difficult to calibrate as a result of their intractability. Accurate prediction of behavior is also a challenge: among the game-theoretic studies of airline competition that do attempt to validate equilibrium results and empirical behavior, predictions often significantly diverge from observed behavior (e.g., [4]). We address many of these challenges in this research project.

III. MODEL

As mentioned in Section II, our two-stage frequency-fare model is most consistent with the actual airline decisionmaking behavior. In this model, the frequency decisions of all airlines in all nonstop segments are assumed to be made in the first stage, while the average fare decisions for all airlines in all markets are assumed to be made in the second stage. A Subgame-Perfect Nash Equilibrium (SPNE) is the most commonly used solution concept for solving such two-stage games. The SPNE solution concept, in the context of our twostage frequency-fare game, states that for any given set of frequency decisions of all airlines, the fares are modeled in the second stage using the classical Nash equilibrium concept, which states that each airline sets its own fares to maximize its own profit. Next, building on this idea, SPNE concept dictates that the first-stage frequency decisions of each airline are made to maximize that airline's own profit, while explicitly accounting for the corresponding fare decisions as dictated by the second-stage fare equilibrium.

Our model uses the following notation. We define a market as an origin-destination pair of airports, and denote by K_a the set of markets in which an airline α competes. The set of competing airlines is denoted by A . Revenue for an airline α in market m is computed as

$$
Re v_{a,m} = \min(M_m * MS_{a,m}, f_{a,m} * s_{a,m}) * p_{a,m} \quad (1)
$$

where M_m is the size of the market m . $s_{a,m}$ is the seating capacity per flight, and $MS_{a,m}$ is the market share, for airline a in market m . We model market share using a multinomial logit model, an approach widely used in prior literature on air travel demand modeling [21]. We use two alternative utility specifications in terms of the relationship between utility and frequency, one based on the often-cited "S-curve" relationship between frequency share and market share, and the other based on the concept of schedule delay, defined as the difference between the actual flight departure time and passengers' most desired departure times. In the former, utility is given as a linear combination of fare and a logarithmic transformation of frequency, consistent with the S-curve model of the relationship between market share and frequency share ([10], [13]). Let the set of airlines competing in market m be A_m . Let positive parameters α and β respectively indicate passengers' sensitivity to frequency and fare changes. Let N_m be the exponential of the utility of the no-fly alternative. Then the market share of airline a in market m can be expressed as:

$$
MS_{a,m} = \frac{\exp(\alpha \ln(f_{a,m}) - \beta p_{a,m})}{N_m + \sum_{i \in A_m} \exp(\alpha \ln(f_{i,m}) - \beta p_{i,m})}
$$
(2)

With equal fares and in the absence of a no-fly alternative, market share in (2) is a function of only the frequency share, following an S-curve whose shape is modulated by α . Market share can also be captured using the schedule-delay model, as discussed in $[13]$. In this case, market share for airline α takes the following form:

$$
MS_{a,m} = \frac{\exp(-\varphi_{a,m} - r - \beta p_{a,m})}{N_m + \sum_{i \in A_m} \exp(-\varphi_{i,m} - r - \beta p_{i,m})}
$$
(3)

Here φ and r are positive parameters modulating the utility of frequency. Hansen and Liu argue that this model describes a more plausible relationship between frequency share and market share [13]. For instance, market share in this model depends on both frequency share and competitor frequency, such that an airline cannot simply dominate the market share of an already high frequency market by arbitrarily increasing its own frequency, unlike in the S-curve formulation.

We assume operating costs to be linear in frequency, i.e., $Cost_{a,m} = c_{a,m} * f_{a,m}$ for airline *a* and market *m*, where $c_{a,m}$ is the cost per flight for airline a in market m . The overall payoff (or profit) function of airline a is then given by:

$$
\pi_a = \sum_{m \in K_a} \left(Rev_{a,m} - Cost_{a,m} \right) \tag{4}
$$

IV. ANALYTICAL RESULTS

Using the airline payoff function given by (4), we first analyzed a simplified version of our model, with a single market, two airlines, no connecting passengers, unlimited seating capacity, and the absence of no-fly alternative (i.e., all passengers in the market have to choose one of the two airlines). The purpose of analyzing this simplified model was to generate insights into the analytical and tractability issues underlying this model formulation. For this simplified model, we proved the following three analytical properties, for both the S-curve model and the schedule delay model. In this paper, we simply state them without proving. Interested readers are referred to [3] for more details and proofs.

Mathematical Property 1: The second-stage fare game always has a unique pure strategy Nash equilibrium.

Mathematical Property 2: For the first stage game, each airline's payoff is concave in its own frequency strategy. That is, the second derivative of each airline's payoff in its own frequency is negative.

Mathematical Property 3: For the first stage game, each airline's profit function is submodular in the overall frequency space. That is, the cross derivative of each airline's payoff, with respect to the frequencies of both airlines is negative.

These results are significant because they demonstrate that subgame-perfect Nash equilibrium is a credible and tractable solution concept for our simplified two-stage game. In particular, the existence and uniqueness results indicate the suitability of using pure strategy Nash equilibrium as a solution concept for the second-stage game. Submodularity of the first-stage game ensures that the existence of a purestrategy Nash equilibrium is guaranteed, and a broad class of adaptive learning dynamics (including successive adaptive learning dynamics (including successive optimizations algorithm, and fictitious play) converge to the interval bounded by the Nash Equilibria with the largest and smallest frequencies ([22], [23]). If there is a unique equilibrium, these dynamics converge to it. Concave firststage payoffs, as proved by us here, are not guaranteed for one-stage frequency models ([4]). For our two-stage model, however, they ensure that first stage payoff maximization problems (as part of a successive optimizations algorithm, for instance) are efficiently solvable and have a unique optimum.

This analysis implies that a two-stage game approach to modelling frequency and fare competition induces properties in the payoff functions that improve the credibility and tractability of subgame perfect Nash equilibrium. In other words, a more realistic approach to the modeling of the sequential nature of airline decision-making makes the corresponding game-theoretic model more attractive, both computationally and behaviorally. The existence of these properties in this simple case suggests that more complex models may also show some similar favorable properties. However, analytical approaches become substantially more difficult when the strong assumptions of this simplified form of our model are progressively relaxed. Therefore, we turn to numerical and computational approaches to extend our results to more realistic models. The following section describes these approaches and their results.

V. NUMERICAL EXTENSIONS

In this section, we present results of a series of numerical experiments for solving the more realistic full version of our model (as against the simplified version solved analytically). We now relax the five assumptions (namely, single market, two airlines, no connecting passengers, no no-fly alternatives, and unlimited seats) of our simplified model one by one, and numerically test the existence, uniqueness, concavity and submodularity results for a range of parameter values. In order to do this, we use polynomial approximations of second-stage payoffs as functions of frequencies of all airlines in that market. These functions provide a very good fit to the exact payoff functions, and allow for convenient evaluation of game properties. Additionally, in the real-world, airlines can make frequency decisions with only an approximate knowledge of what the likely fare levels will be in the future. Therefore, a reasonable approximation of payoff functions that captures the gross properties of these functions seems justified.

A. Solving the Second-Stage Game

We first compute equilibrium fare solutions of the secondstage game for any given combination of plausible first-stage frequency decisions. Second-stage equilibria were computed by initializing fares for all players (i.e., airlines) at \$100 (arbitrarily), and numerically optimizing each player's payoff (given by (4)) one by one with respect to that player's fare. This iterative optimization algorithm was repeated until fares for all player converged to within a small threshold (that is, within a change of less than \$0.1 from the previous iteration). This was done for both S-curve and schedule delay market share functions; for 1, 2 and 3 player markets; for varying numbers of seats-per-flight; with and without connecting passengers; for varying values of the exponential of the utility of the no-fly alternative (i.e., for varying values of N_m); and for varying values of the utility parameters for frequency and fare (i.e., for varying values of β and α for the S-curve model, and φ and τ for the schedule-delay model). Ranges of varied parameters were chosen to encompass values found in literature and in practice. Tables I and II list the ranges tested for each parameter and the increments in which these parameters were varied for the S-curve and schedule-delay models respectively.

TABLE I. PARAMETER RANGES AND INCREMENTS FOR THE S-CURVE MODEL

Parameter	Range Tested	Testing Increments	
	0 to 1	0.1	
α	1 to 2	0 ₁	
	0.001 to 0.01	0.001	
Seats-per-flight (S)	25-250, and unlimited seating	25	

TABLE II. PARAMETER RANGES AND INCREMENTS FOR THE SCHEDULE DELAY MODEL

For these numerical experiments, market size M_m was set at 1000 passengers and operating cost was set at \$10,000 per flight. For computational feasibility, parameters were varied one at a time for each of the cases with 1, 2 and 3 players. The following default values of parameters, all set according to the prevailing estimates in the existing literature, were used: $\alpha = 1.29$, $\beta = 0.0045$, and $N_m = 0.5$ for the S-curve model, and $r = 0.456$, $\varphi = 5.1$, $\beta = 0.012$ (as per [13]), $N_m =$ 0.005. The default number of seats was set to 125 for both models. In all cases, this successive optimizations algorithm converged to an equilibrium, suggesting that second-stage fare equilibria for our model exist in practice across a broad range of scenarios. These results are consistent with the analytically demonstrated existence and uniqueness result of the Mathematical Property 1 even for a broader landscape of more realistic but analytically intractable scenarios.

B. Polynomial Approximations of Payoff Functions

For each frequency combination and parameter combination, we recorded equilibrium payoffs π_i for $i \in A$. For each player, we assumed that the plausible frequency values are all integers between 0 and 20. Then, for each parameter combination and for each player, this generated 20 payoff data points for a monopolistic market, 400 data points for a two-player market, and 8000 data points for a threeplayer market. We then fit polynomial functions of frequency values to each payoff function using simple linear regression. Polynomial (specifically quadratic) payoff coefficients were estimated for the following functional forms. For a monopolistic market, the profit of airline 1, π_1 , operating a daily frequency of f_1 was modeled as follows:

$$
\pi_1 \sim \gamma_0^+ \ \gamma_1 f_1^+ + \gamma_2 f_1^2 \tag{5}
$$

For two-player markets, the profit of airline 1, π_1 , operating a daily frequency of f_1 against a competitor operating a daily frequency f_2 , was modeled as follows:

$$
\pi_1 \sim \gamma_0 + \gamma_1 f_1 + \gamma_2 f_2 + \gamma_3 f_1^2 + \gamma_4 f_2^2 + \gamma_5 f_1 f_2 \tag{6}
$$

Polynomial payoff approximations for markets with more than two players were constructed similarly. Table III gives an illustrative example of regression results for a two-player nonstop S-curve model. In this case, utility parameters are held at defaults ($\alpha = 1.29$, $\beta = 0.0045$, $N_m = 0.5$) and the number of seats-per-flight are varied to include various aircraft sizes.

TABLE III. EXAMPLE REGRESSION COEFFICIENTS AND MODEL R^2 for TWO-PLAYER NON-STOP S-CURVE MODEL

S	Yo	γ,	γ,	γ_3	Y4	γ_{5}	\mathbf{R}^2
250	122200	18135	-17856	-494	686	-533	0.96
225	122250	18130	-17861	-494	686	-532	0.96
200	122400	18115	-17876	-493	686	-532	0.96
175	122640	18095	-17901	-493	687	-531	0.96
150	123470	18030	-17989	-492	689	-529	0.96
125	125340	17925	-18214	-490	696	-523	0.95
100	129430	17885	-18838	-495	716	-514	0.94
75	136710	18277	-20301	-512	773	-518	0.93
50	142620	20224	-22355	-567	865	-578	0.93
25	104880	27814	-18929	-743	773	-846	0.97

The coefficient of determination (R^2) for fitted models remained > 0.9 for nearly all tested parameter combinations, ranging from close to 0.9 on the low end (in three-player games with 25 seats per flight, an extreme parameter value) to very close to 1. Fig. 1 shows an example quadratic fit to second stage profits values in a two-player game following the S-curve market share formulation with α at 1.29, β at 0.0045, N_m at 0.5, and 125 seats.

Exceptions were found in extreme or nonsensical parameter combinations: for example, in one-player markets with the S-curve model and very high β (0.009 or 0.01), \mathbb{R}^2 dipped to 0.88 and 0.87 respectively, while in one-player and three-player markets with the S-curve model and $N_m = 0$ (i.e., in the absence of the no-fly alternative), R^2 fell to 0.08 and 0.89 respectively. The uniquely poor fit found in the monopolistic markets with no no-fly alternative is not surprising, because in such markets an airline unrealistically has strong incentive to charge extremely high ticket prices. The generally found high R^2 values suggest that in nearly all cases, a quadratic function of player frequencies is able to capture a significant portion of the variation in equilibrium profits, and can provide a good numerical approximation of the payoff functions described in (4), irrespective of whether we assume an S-curve model or a schedule delay model. This gives us a tool to probe the robustness of the concavity and submodularity properties described in Mathematical Property 2 and 3.

Examining quadratic approximated payoff functions, we find that in all cases with high R^2 (>0.9), the signs of estimated coefficients are consistent with both submodularity and concavity properties. For example, for the two-player case, γ_3 , the coefficient of the square of player 1's daily frequency, and γ_5 , the coefficient of the interaction term $f_1 f_2$, are both negative, consistent with the concavity and submodularity properties respectively. Note that this is the case across the range of all seat values in Table III.

Figure 1. Player 1 payoffs at fare equilibria for various frequency combinations in a two-player game, with S-curve model, $\alpha = 1.29$, $\beta =$ 0.0045, $N_m = 0.5$ and 125 seats per aircraft

While longer computational times precluded extensive parameter sensitivity tests for more than three players, some limited testing of four-player games revealed similar results: good quadratic function approximations with high \mathbb{R}^2 , and coefficient estimates consistent with concavity and submodularity properties. We also examined higher order polynomial approximations for even closer fits to payoff functions: quartic approximations tested on several models retained submodularity and concavity properties. For the remainder of this paper, however, we will focus on quadratic approximations, as these allow for generally good approximations while remaining convenient for quick evaluation of function properties and keeping the number of parameters in check when calibrating models with real-world data.

C. Game Dynamics and Convergence Properties with Approximated Payoff Functions

The robustness of submodularity and concavity properties in approximated payoff functions across a wide range of scenarios and parameter values extends the analytical results of Mathematical Properties 1 and 2 to a much richer and more realistic class of models. These results suggest that in general, sub-game perfect Nash equilibrium remains a highly tractable and credible solution concept for our game. That the property of submodularity extends to more complex scenarios is consistent with the observation that games with this property tend to arise in strategic situations where there is competition for a resource [24]. In case of our problem of airline competition, that resource is market share. While beyond twoplayer games, successive optimizations algorithm convergence cannot be implied simply by submodularity, analogy with the two-player case, as well as a growing body of literature on games of strategic substitutes (e.g., [24], [25]) provide us with a baseline for further exploration of the convergence properties for submodular games with more than two players.

Concave payoffs still ensure existence of first-stage frequency equilibrium, and our quadratic approximations provide a simple mechanism for checking the uniqueness of first-stage equilibrium using Rosen's diagonal strict concavity condition [26]. We find that with a few exceptions in extreme parameter values, e.g., high α values (> 1.7), estimated coefficients across our tested parameter ranges are consistent with guaranteed unique first-stage equilibrium. Furthermore, concavity means that individual players' optimization problems give unique optimal solutions, and can be easily solved to optimality. Thus, the successive optimizations algorithm can be deployed efficiently to find equilibria, even in large-scale scenarios and networks. Thus, on the one hand, for the two-player case, we can use submodularity and the uniqueness of first stage Nash equilibrium to guarantee the rapid convergence of a broad class of adaptive dynamics to Nash equilibria in the first-stage game (following [2]) with our approximated payoff functions. On the other hand, for the games with more than two players, we can use concavity, submodularity, the polynomial nature of approximated payoffs, and results from Jensen [25] to demonstrate the convergence of successive optimizations algorithms.

The convergence of successive optimizations algorithm in games with our approximated payoff functions is reassuring

both from an intuitive and a computational perspective. It allows us to model, using the concept of Nash equilibrium, even those situations where the airlines in our game make decisions in a less-than perfectly rational way. Even if the airlines are not deemed to possess the infinite rationality that generally underlies the Nash equilibrium solution concept, the convergence property of the myopic approach involving successive optimizations ensures that the Nash equilibrium is reached even with myopically rational players. From a computational point of view, fast convergence of easily implementable successive optimizations algorithm and easily solvable individual payoff maximizations enable efficient solutions, experimentation, and calibration of our model when comparing its predictions to observed airline behavior. In the next section, we exploit these properties to apply our model to a real-world airline network.

VI. AIRLINE NETWORK CASE STUDY

To test the tractability and forecasting accuracy of our model in practice, we apply our game-theoretic model to a network of airports in the western United States. The test network consists of 11 airports, namely, Seattle-Tacoma International Airport (SEA), Portland International Airport (PDX), San Francisco International Airport (SFO), San Diego International Airport (SAN), Los Angeles International Airport (LAX), Las Vegas McCarran International Airport (LAS), Phoenix Sky Harbor International Airport (PHX), Oakland International Airport (OAK), Ontario California International Airport (ONT), Sacramento International Airport (SMF), and Mineta San Jose International Airport (SJC). Our datasets spans across the eight year period of 2007-2014. We estimate the daily non-stop flight frequencies of the four major airlines in this network during this period, namely, Alaska Airlines (AS), United Airlines (UA), US Airways (US), and Southwest Airlines (WN) in the markets in which they are present by computing the first-stage Nash equilibrium using concave, submodular quadratic functions to approximate the payoffs.

A. Data Sources

Quadratic payoff functions were constructed for each valid airline-market combination depending on the number of airlines in the market, based on actual cost and market size data taken from the Bureau of Transportation Statistics (BTS) records. Payoff function approximations computed for default parameters and market sizes were transformed by the operating costs and demands in each particular market, a simple transformation given the functional form of (4). Operating costs and airborne hours for different aircraft for different airlines was obtained from the Schedule P-5.2 tables from the BTS website [27]. Data on market size, observed frequencies and flight distances was obtained from the T100 Segments tables on the BTS website [28]. Data from unidirectional markets containing the same airports were averaged, such that, for example, PDX-SAN and SAN-PDX were treated identically for payoff function generation and frequency estimation purposes, as passenger flows, observed frequencies, and other data were generally quite similar for differently ordered airport pairs. For simplicity, airline-market combinations with an airline's market-share of less than 10%

or average daily frequency of less than 0.5 were removed from consideration.

B. Network Payoff Functions

We used the successive optimizations algorithm, justified analytically and numerically by the results in Sections IV and V, to solve the polynomial approximation of the first-stage game. Within each individual optimization, an airline decides its vector of frequencies over all its nonstop segments to maximize the sum of the payoff functions across all its markets. Given the large proportion of nonstop passengers in our dataset, we assumed all markets to be nonstop, and simply used segment passenger flows as market demands for this case study. The iterative algorithm was run until estimated frequencies of all airlines converged to within a tolerance threshold. We constrained the feasible frequency decisions by the estimated availability of various aircraft types to the airline within the network. To estimate aircraft availability within the network, airlines were assumed to generally utilize aircraft close to the limits of availability. Thus, the number of aircraft of type k available to a certain airline α was calculated as:

$$
F_{k,a} = \frac{\sum_{m \in M_{k,a}} 2*f_{k,a,m}*(b_{a,m}+t)}{T}
$$
 (7)

Here, $M_{k,a}$ the set of segments where airline α uses aircraft type k , $f_{k,a,m}$ is the observed frequency of airline a in segment m using aircraft type k , t is the turnaround time of the aircraft (taken to be 30 minutes in all cases), $b_{a,m}$ is the average number of block hours per flight of airline α on segment m . the factor of 2 accounts for the two directed segments corresponding to a particular airport pair, and T is the number of hours of available flying time in the day, which we assume to be 18 on average. During each individual airline profit maximization problem, these fleet size restrictions were applied such that:

$$
\sum_{m \in M_{k,a}} 2 * \hat{f}_{k,a,m} * (b_{a,m} + t) \le T.F_{k,a}
$$
 (8)

In (8), $\hat{f}_{k,a,m}$ is the model estimated frequency on segment m for airline a using aircraft type k . In all, our network had 68 combinations of airlines and segments for which the frequency estimation was conducted. When the successive optimizations algorithm is run, the players are assumed to allocate flight frequencies across their respective networks by solving a constrained quadratic optimization problem during each iteration, continuing until convergence. Frequency decision vectors for each player were initialized at 0. The model typically converged in 6-7 iterations (with each iteration consisting of four optimizations, one for each airline), and is solved in less than one second using MATLAB quadratic programming functions.

C. Payoff Calibration and In-Sample Performance

In order to calibrate our model, we adjust payoff polynomial coefficients to minimize the Mean Absolute Percentage Error, or MAPE, between estimated and observed frequencies over the networks of all airlines. MAPE is calculated as:

$$
MAPE = \frac{\sum_{as \in AS} |\hat{f}_{as} - f_{as}|}{\sum_{as \in AS} f_{as}} \tag{9}
$$

Here, AS is the set of all airline-segment combinations. \hat{f}_{as} is the estimated frequency, and f_{as} is the observed frequency, for the airline-segment combination as . For the purposes of calibration, airline-segment combinations were divided into four groups: three-player markets, two-player markets where both airports were hubs for the airline, other two-player markets, and monopolistic markets. The resulting 11 payoff function coefficients (the linear, quadratic, and interaction term coefficients) were adjusted simultaneously before transformation by cost and market size data for each airlinesegment combination.

These coefficients were adjusted using a gradient approximation algorithm called SPSA (Simultaneous Perturbation Stochastic Approximation, from [29]) to minimize overall MAPE. Specifically, during each iteration of SPSA, a single game was solved, with payoff coefficients perturbed according to an approximated gradient with respect to the MAPE loss function. SPSA was chosen for its ability to approximate gradient using only two measurements of the $MAPE$ loss function, independent of the number of variables being optimized. The 11 coefficients were initialized using values estimated by fitting quadratic functions of frequency to payoff using the S-curve market share model with $\alpha = 1.29$, $\beta = 0.0045$, $N_m = 0.5$ and unlimited seating. The game was then run repeatedly until approximate convergence of $MAPE$, over the course of roughly 10,000 iterations. The best performing coefficients were then used to estimate frequencies across the network, from which we can evaluate in-sample and out-of-sample model performance.

First, we used the data from the first quarter of 2007 to calibrate the model and then compared the calibrated model's MAPE performance against the actual frequency values for the same quarter. This gave us an estimate of the in-sample prediction accuracy of our model. Fig. 2 compares actual frequencies (x-axis) and these predicted frequencies (y-axis). The 45º blue line represents perfect predictions, i.e., the line where the observed and predicted frequencies are equal. Most data points indicating segment predictions are near this line. An overall in-sample MAPE of 18.4% is achieved: more concretely, this corresponds to 49% of absolute prediction errors being less than 1, and 78% being less than 2. Notable outliers were the three highest frequency segments, all hub-tohub airport segments flown by Southwest Airlines (circled in Fig. 2).

Figure 2. Actual versus model-predicted frequency, for Q1 of 2007

D. Out-of-Sample Predictive Performance

We can use these same calibrated coefficients to then make out-of-sample frequency predictions for future quarters. We call the dataset used for calibration of the model coefficients as the training dataset and the dataset used for testing the prediction accuracy as the testing dataset. For example, coefficients calibrated using SPSA on data from Q1 of 2007 (the training dataset) can be used to predict frequencies for Q4 of 2007 (the testing dataset). For this training-testing pair, we found an out-of-sample testing $MAPE$ of 20.6% , corresponding to 47% of absolute frequency errors being less than 1, and 73% being less than 2. However, we can leverage our knowledge of training errors when making out-of-sample predictions to further improve this performance. By adjusting our testing predictions for a given airline-segment by the error for that same airline-segment in the training dataset (and simply performing no adjustment to airline-segments that did not exist in the training dataset), we can substantially reduce our out-of-sample $MAPE$: $MAPE$ in our Q4 2007 predictions falls to just 11.2%, corresponding to 72% of absolute frequency errors being less than 1, and 92% being less than 2.

A more concrete illustration of this model's out-of-sample prediction accuracy can be found by looking at a new market that arises between Q1 and Q4 of 2007. PDX-SFO is not seen in the training data, yet in Q4 it is a duopoly market shared by AS and UA. The frequencies predicted for this market are a good approximation for observed behavior, with an overall $MAPE$ of 16.5%, as seen in Table IV.

We can take a another view of out-of-sample prediction accuracy by looking at more aggregate measures of prediction performance, at the airline, coefficient category (one-player, two-player hub-hub, other two-player, three-player), market and airport levels. In our Q4 2007 predictions based on Q1 calibration data, we find excellent predictions at all of these levels, both unadjusted and adjusted according to training error as described above. With respect to total frequencies allocated by each airline, we find an $MAPE$ of 2.0%, or 1.5% adjusted (corresponding to average absolute errors of 2.71 and 2.11 flights, respectively). With respect to total frequencies allocated within each coefficient category, we find an MAPE of 3.0%, or 2.5% adjusted (corresponding to average absolute errors of 4.2 and 3.42 flights, respectively). With respect to total frequency by market (across the 41 markets for which estimates were made in the network), we find an MAPE of 14.4%, or 6.3% adjusted (corresponding to average absolute errors of 1.95 and 0.86 flights, respectively). With respect to total frequency by airport (across the 11 airports in the network), we find an MAPE of 7.8%, or 2.6% adjusted (corresponding to average absolute errors of 7.75 and 2.59 flights, respectively). Predictions at each of these levels of aggregation may be of interest for airlines, airports and other policy makers.

TABLE IV. OUT-OF-SAMPLE PERFORMANCE ON A NEW MARKET

PDX-SFO, 04 2007 coefficients With calibrated on O1 2007	Observed Frequency	Predicted Frequency	Absolute Error
UA	6.11	7.34	.22
AS	3.02	2.74	0.28

Table V displays the (rounded) error adjusted predictions and actual total daily flights for each airport in our network for Q4 2007, while Fig. 3 shows a visual representation of this table. Note that at airports where a significant increase in traffic levels between Q1 and Q4 was observed (e.g. LAS), our model was able to predict this increase. As discussed in Section I, these types of airport-level traffic forecasts are used, for instance, by the FAA in workforce staff planning, and evaluation of airport capacity expansion and technological development.

In order to examine the predictive accuracy of our model more broadly, we can examine in-sample and out-of-sample prediction across years and for varying degrees of look-ahead in prediction. In order to do this, we calibrated our 11 coefficients on every quarter from 2007 to 2014, giving us 32 sets of coefficients. Then, we used these coefficients to predict frequencies for every quarter after each of these calibration quarters. In this expanded set of data, we include new major airlines and new hubs in the network as appropriate for the time period in question. Examining the adjusted $MAPE$ at varying look-ahead values (i.e., at varying differences between the quarters corresponding to the training and testing datasets), we find an almost monotonic increase in median error.

TABLE V. OUT-OF-SAMPLE PREDICTIONS OF DAILY FLIGHTS AT AN AIRPORT LEVEL: Q1 2007 FOR TRAINING AND Q4 2007 FOR TESTING

Airport	Observed Flights Q1	Observed Flights Q4	Predicted Flights Q4
LAX	137	128	132
SJC	60	64	61
LAS	154	167	168
SAN	101	108	110
SMF	70	71	70
SEA	98	105	104
PDX	33	41	43
SFO	67	92	99
ONT	61	61	62
PHX	159	159	153
OAK	88	88	85

Figure 3. Visual representation of Table V data. Left panel shows airports where flight numbers remained roughly the same, the middle panel airports where flights decreased, and the right airports where flights increased.

Fig. 4 shows MAPE for each possible training-testing combination (red circles) in our 8-year dataset, and the median MAPE for that look-ahead value in blue, with error bars at one standard deviation. Fig. 5 shows the same information as that in Fig. 4, but focuses only on new markets, which did not exist in the training dataset but do exist in the testing dataset. Fig. 6 is similar, but shows aggregate MAPE values calculated on airport level. Median MAPE remains below 15% across all markets, remains around 20% for the new markets, and remains below 10% for the airport level aggregation, for several quarters out, suggesting reasonable predictive accuracy in the short and medium terms.

VII. CONCLUSIONS

This study investigates a two-stage frequency-fare gametheoretic model of airline competition which is behaviorally consistent with the sequential nature of airline frequency and fare decisions. For simple cases, the analytical qualities of this model indicate well behaved and tractable games, with unique equilibria and convergence properties. Using polynomial payoff function approximations, these properties can be shown numerically to extend to more realistic formulations of the game. In practice, when applied to a real airline network, the model converges quickly and generates daily frequency predictions that closely approximate actual airline decisions, both in-sample and out-of-sample, within a short-to-medium term time horizon.

We believe that our model presents multiple avenues for application and future research within air traffic forecasting and related decision-making. To the best of our knowledge, this is the first study to investigate the favorable properties discussed within the context of a two-stage model of airline competition, providing analytical, computational, and empirical results for a game-theoretic approach that has received limited attention in the airline competition literature. We hope that our results presented here can serve as a foundation for a further research into sequential models of airline decision-making under competition. Furthermore, the predictive performance of our model on real world data suggests that refinements of the model could serve as a scenario analysis tool to aid in planning, forecasting and policy-making decision support. The tractability of this model and the flexibility with which different scenarios can be tested suggest its potential for rapid and interpretable experimentation in even large-scale airline networks. Here we have considered a relatively simple model of airline competition, without taking into account factors such as market segmentation between business and leisure passengers, passenger loyalty, behavioral differences between airlines and between the other amenities that they provide, and characteristics of markets beyond cost, observed passenger flow, and hub presence. The fact that our simple model provides a good approximation of airline frequency allocation suggests that more flexible parameterizations in calibration and prediction could be promising avenues for practitioners.

Figure 4. MAPE values (adjusted for calibration error) for varying lookahead durations for all markets

Figure 5. MAPE values (adjusted for calibration error) for varying lookahead durations for new markets

Figure 6. Airport level aggregated MAPE values (adjusted for calibration error) for varying look-ahead durations

REFERENCES

- [1] Federal Aviation Administration (FAA), "Terminal area forecastmodernization (TAF-M) methodology," Office of Aviation Policy and Plans, APO-100, Dec 2016.
- [2] D. Bhadra, J. Gentry, B. Hogan, and M. Wells, "Future air traffic timetable estimator," Journal of Aircraft, vol. 42, no. 2, pp. 320–328, March-April 2005.
- [3] R. Harder, and V. Vaze, "Two stage game theoretic modeling of airline frequency and fare competition," unpublished.
- [4] M. Hansen, "Airline competition in a hub-dominated environment: an application of non-cooperative game theory," Transportation Research Part B: Methodological, vol. 24, no. 1, pp. 27–43, February 1990.
- [5] N., Adler, "Competition in a deregulated air transportation market," European Journal of Operational Research, vol. 129, no. 2, pp. 337–345, March 2011.
- [6] N. Adler, "Hub-spoke network choice under competition with an application to western Europe," Transportation Science, vol. 39, no. 1, pp. 58–72, February 2005.
- [7] N. Adler, E. Pels, and C. Nash, "High-speed rail and air transport competition: game engineering as tool for cost-benefit analysis," Transportation Research Part B: Methodological, vol. 44, no. 7, pp. 812– 833, August 2010.
- [8] V. Vaze, and C. Barnhart, "Modeling airline frequency competition for airport congestion mitigation," Transportation Science, vol. 46, no. 4, pp. 512–535, May 2012.
- [9] V. Vaze, and C. Barnhart, "Airline frequency competition in airport congestion pricing," Transportation Research Record, vol. 2226, pp. 69– 77, 2012.
- [10] V. Vaze, and C. Barnhart, "Price of airline frequency competition," in Game-Theoretic Analysis of Congestion, Safety and Security, K. Hausken and J. Zhuang, Eds. Springer Series in Reliability Engineering, 2015, pp. 173–217, ISBN: 978–3–319–13008–8.
- [11] S. Hong, and P. T. Harker, "Air traffic network equilibrium: toward frequency, price and slot priority analysis," Transportation Research Part B: Methodological, vol. 26, no. 4, pp. 307–323, August 1992.
- [12] P. Zito, G. Salvo, and L. La Franca, "Modelling airlines competition on fares and frequencies of service by bi-level optimization," Procedia Social and Behavioral Sciences, vol. 20, pp. 1080-1089, 2011.
- [13] M., Hansen, and Y. Liu, "Airline competition and market frequency: a comparison of the S-curve and schedule delay models," Transportation Research Part B: Methodological, vol. 78, pp. 301–317, August 2015.
- [14] W. Wei, and M. Hansen, "Airlines' competition in aircraft size and service frequency in duopoly markets," Transportation Research Part E: Logistics and Transportation Review, vol. 43, no. 4, pp. 409–424, July 2007.
- [15] J. K. Brueckner, "Schedule competition revisited," Journal of Transport Economic Policy, vol. 44, no. 3, pp. 261–285, September 2010.
- [16] G. Dobson, and P. J. Lederer, "Airline scheduling and routing in a huband-spoke system," Transportation Science, vol. 27, no. 3, pp. 281–297, August 1993.
- [17] Y. Schipper, P. Rietveld, and P. Nijkamp, "Airline deregulation and external costs: a welfare analysis," Transportation Research Part B: Methodological, vol. 37, no. 8, pp. 699–718, September 2003.
- [18] J. K. Brueckner, and R. Flores-Fillol, "Airline schedule competition," Review of Industrial Organization, vol. 30, no. 3, pp. 161–177, May 2007.
- [19] V. D. Norman, and S. P. Strandenes, "Deregulation of scandinavian airlines: a case study of the Oslo-Stockholm route," in Empirical Studies of Strategic Trade Policy, P. Krugman, A. Smith, Eds. University of Chicago Press, Chicago, 1994, pp. 85–100.
- [20] V. Aguirregabiria, and C. Y. Ho, "A dynamic oligopoly game of the US airline industry: estimation and policy experiments," Journal of Econometrics, vol. 168, no. 1, pp. 156–173, May 2012.
- [21] L. A. Garrow, Discrete Choice Modelling and Air Travel Demand Theory and Applications. Surrey, United Kingdom: Ashgate Publishing, 2012, ISBN: 978-1-4094-8633-6.
- [22] P. Milgrom, and J. Roberts, "Rationalizability, learning, and equilibrium in games with strategic complementarities," Econometrica, vol. 58, no. 6, pp. 1255-1277, November 1990.
- [23] Y. Chen, and R. Gazzale, "When does learning in games generate convergence to Nash equilibria? The role of supermodularity in an experimental setting," American Economic Review, vol. 94, no. 5, pp. 1505-1535, December 2004.
- [24] S. Roy, and T. Sabarwal, "Characterizing stability properties in games with strategic substitutes," Games and Economic Behavior, vol. 75, no. 1, pp. 337-353, May 2012.
- [25] M. K. Jensen, "Aggregative games and best-reply potentials," Economic Theory, vol. 43, no. 1, pp. 45–66, April 2010.
- [26] J. B. Rosen, "Existence and uniqueness of equilibrium points for concave N-person games," Econometrica, vol. 33, no. 3, pp. 520-534, July 1965.
- [27] Bureau of Transportation Statistics, Air Carrier Financial: Schedule P-5.2, Office of the Assistant Secretary for Research and Technology, Bureau of Transportation Statistics, United States Department of Transportation, Washington DC, 2016.
- [28] Bureau of Transportation Statistics, Air Carrier Statistics Database (T100), Office of the Assistant Secretary for Research and Technology, Bureau of Transportation Statistics, United States Department of Transportation, Washington DC, 2016.
- [29] J. Spall, "An overview of the simultaneous perturbation method for efficient optimization," John Hopkins Technical Digest, vol. 19, no. 4, pp. 482-492, 1998.

AUTHOR BIOGRAPHY

Vikrant Vaze is an Assistant Professor at the Thayer School of Engineering at Dartmouth College. His research focuses on improving the planning, management and operations of large-scale, multi-stakeholder systems such as transportation and healthcare using game theory, optimization and data analytics. He is the recipient of several awards, including the Milton Pikarsky Best PhD Dissertation Award from CUTC, the Anna Valicek Best paper Honorable Mention Award from AGIFORS, Best Paper Award from the ATM Seminar, President of India Gold Medal, MIT Presidential Fellowship, UPS Fellowship, Airport Cooperative Research Program Scholarship from the US DOT, and the National Talent Scholarship from the Government of India. He received an MS in Transportation, another MS in Operations Research and a PhD in Systems, all from the Massachusetts Institute of Technology.

Reed Harder is a PhD candidate in Engineering Sciences at the Thayer School of Engineering at Dartmouth College. He received Bachelor's degrees in Engineering and History from Dartmouth College in 2013. His research interests include game theory, statistical inference, and optimization, with particular application to multi-agent systems in transportation and healthcare. He has also done work in statistical genetics and citation network analysis. He has presented his research at the 11th Workshop on the Economics of Networks, Systems and Computation in Juan-les-Pins, France and the INFORMS 2016 Annual Meeting in Nashville, TN, where he received an Honorable Mention Award for Best Student Presentation in Aviation Applications.