Structural Forecasting of Arrivals to Scheduled Services

Fernando Bernstein[1] N. Bora Keskin[1] Adam Mersereau[2] Morgan Wood[3] Serhan Ziya[3]

[1]Fuqua School of Business, Duke University

[2]Kenan-Flagler School of Business, University of North Carolina at Chapel Hill

[3]Statistics & Operations Research Department, University of North Carolina at Chapel Hill

Motivated by passenger arrivals to the security checkpoint at the Raleigh-Durham International Airport, we develop methods to forecast future arrivals to a system in which arrivals are tied to scheduled events. We consider a setting where the historical number of passenger arrivals is known, but it is unknown for which flight each passenger arrived. In this setting, we propose methods to estimate a parametric distribution of passenger arrival earliness, called a "show-up profile." Estimated show-up profiles, along with future flight schedules, are used to forecast numbers of future passenger arrivals. We explore the benefit of allowing the show-up profile to vary depending on time of day, day of week, and flight-type. We compare the performance of these passenger arrival estimates to estimates derived from machine learning methods using data collected at the airport using real-time sensors. We show that machine learning methods and structural methods have similar performances and that structural methods can be improved by considering different show-up profiles for different times and destinations. Compared to machine learning methods, the structural methods allow for better interpretability which we showcase in an empirical analysis of the relationship between passenger earliness and flight distances. Ultimately, combining structural methods with machine learning yields the best forecasting performance.

1. Introduction

A key challenge in running a complex operation like an airport is managing the passenger experience, which is itself a multifaceted task. Passengers are unhappy when they have difficulty finding a parking spot, when they encounter expensive or limited food options, and when they find neglected bathroom facilities. However, a factor that is most associated with passenger dissatisfaction is crowding at the airport. Crowding and delays not only color the passengers' experience while they are physically at the airport, but they may also lead to missed flights which have more serious consequences for passengers. Therefore, airport administrators and designers are deeply interested in managing passenger flows and crowding through reactive measures such as deploying staff for crowd management, short-term controls such as staff scheduling, and long-term decisions such as designing terminals with sufficient capacity. A pre-requisite for managing passenger flows is a model that can predict passenger arrivals (to check-in counters, the security checkpoint line, gates, etc.) and can facilitate sensitivity analysis of different arrival scenarios.

To address this challenge, the industry commonly utilizes "show-up profiles" (also known as "arrival profiles" and "earliness arrival functions"; see Ashford et al. 2011, IATA 1995, Isarsoft 2024). A show-up profile specifies the probability distribution for how early (relative to the flight's departure time) passengers of a flight arrive at a particular location (e.g., the security checkpoint queue) in the airport. The show-up profile for a particular flight allows airlines and airport administrators to anticipate flight-specific crowds at the check-in and boarding areas. More importantly, by stacking the show-up profiles for all flights in a day, one can naturally derive aggregate forecasts of arrivals at commonly used areas such as security checkpoints.

To illustrate the notion and use of a show-up profile, consider the left panel of Figure 1, which plots a show-up profile used by a planner at the Raleigh-Durham International Airport (RDU) in the U.S. state of North Carolina, which we partnered with for this research. Suppose that there are known flight departures at times 9:00, 10:00, and 10:30, each with an equal number of passengers. The right panel in Figure 1 displays the arrival distribution derived from "stacking" the show-up profiles according to scheduled flight departure times for the three flights.

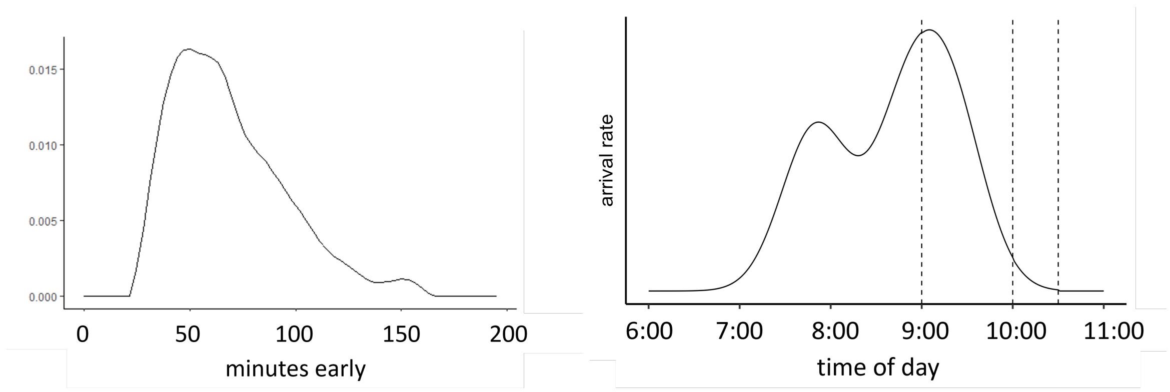

Figure 1 Example of a possible show-up profile—in this figure, an example used by a planner at the RDU airport—and a resulting arrival distribution derived from "stacking" the show-up profile according to scheduled flight departure times which are displayed using dotted vertical lines in the right panel.

Airports are service environments in which customers experience services at scheduled times (i.e., flight departure times) and thus are similar to other service environments such as entertainment venues like movie multiplexes, where customer arrivals are driven by movie start times, or large healthcare facilities, which house multiple clinics and use block scheduling . Therefore, even though this paper is directly motivated by airports, the proposed methodology and results are relevant to a broader class of service environments.

As a basis for forecasts, show-up profiles have several advantages: they are readily understood by managers and they are easily manipulated and used for "what if" analyses. Historically, airport planners have used standard show-up profiles—e.g., a prominent reference manual on airport management (IATA) provides a stock show-up profile that is commonly employed both in practice and academic studies (see Cheng 2014, Ahyudanari and Vandebona 2005)—or they have estimated show-up profiles by surveying arriving customers in airports (see Park and Ahn 2003, Bruno et al. 2019, Rauch and Kljajić 2006). Using standard show-up profiles has obvious shortcomings, as prior work has found that passengers' show-up patterns vary by departure airport and with various factors including destination, time of day, day of week, carrier, and passenger fare class (see Chun and Mak 1999, Postorino et al. 2019). We posit that show-up profiles are also likely to shift seasonally and over time. Therefore, it is valuable to have an inexpensive and automatic way to estimate show-up profiles more dynamically from available data.

Despite the crucial role show-up profiles play in some of the important operational decisions at airports, to the best of our knowledge, no prior work has provided a road map for how they can be estimated in practice and how the estimated show-up profiles can potentially be used for planning purposes. The objective of this paper is to help fill this gap.

Estimating show-up profiles requires data on passenger flows. Recent years have brought identifiable data on passenger arrivals from baggage handling information systems, airline-specific departure control systems, and government security agencies that could aid in estimating show-up profiles (see Postorino et al. 2019). However, baggage handling information systems only reflect passengers checking bags, which is only a (potentially biased) sample of the overall passenger population. Departure control systems can capture passenger check-ins, but the data are owned by airlines and do not capture operationally useful arrival timestamps from passengers checking in online. Finally, government security agencies such as the Transportation Security Agency (TSA) in the United States do not routinely share identified passenger information with airports. This is the case for our airport partner, RDU.

The TSA checkpoint at RDU experiences congestion frequently over the course of a day, serves as a bottleneck in passenger flows, and causes crowding in adjacent areas. This congestion is likely to worsen in the future. RDU, like many airports around the U.S. and globally, has seen record-breaking passenger arrivals in the past year, see Bergin (2024) and Josephs (2024). Moreover, the airport is moving ahead with expansion plans, which include a longer runway to accommodate larger aircraft, and it expects increases in passenger traffic in the coming years. As the airport does not have access to real-time passenger flow data from TSA or the airlines, we installed infrared-beam people counters at the entrances to the TSA security checkpoint area in Terminal 1 of the airport. These sensors report in real time in 15 minute increments the number of passengers arriving at the checkpoint. An example of a day of sensor data can be seen in Figure 2 along with scheduled flight departure times overlaid as vertical lines.

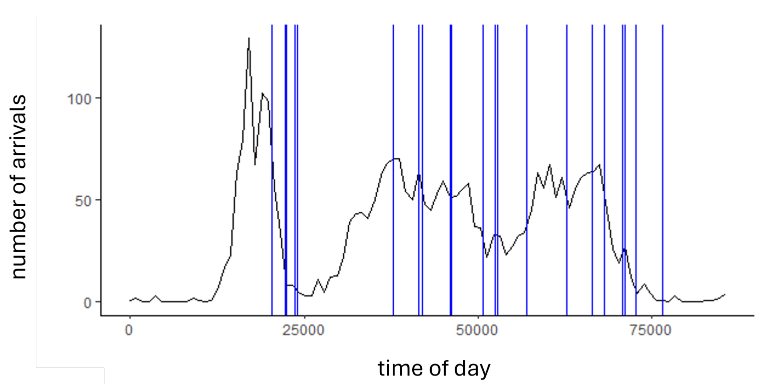

Figure 2 Passenger arrival and scheduled flight departure data from April 15, 2022, at RDU Terminal 1. Sensor counts of the number of arrivals in each 15 minute interval are depicted in black and the scheduled flight departure times are overlaid as vertical lines.

Using sensor data to estimate show-up profiles for individual flights poses significant challenges as it is not possible to match passenger arrivals to flights. (For example, a passenger arriving to TSA at 9:15 am may be 45 minutes early for a 10:00 am flight or 75 minutes early for a 10:30 am flight.) This makes estimation a non-trivial task. If one knew to which flight each passenger tracked by the sensors is assigned, then estimating a show-up profile would be a straightforward density estimation problem; see Postorino et al. (2019). The main challenge we tackle in this work is the estimation of show-up profiles with missing data about the matching of passengers to flights. Another limitation of our data is that it comprises arrival counts aggregated at 15-minute intervals and so is interval-censored.

In this paper, we propose a method for structurally estimating show-up profiles under these data limitations. We consider a parametric model that assumes the show-up profile belongs to a known family of distributions. We assume the number of passengers assigned to each flight is known, but the matching of passengers to flights is not. The problem can be reframed as estimating the parameters of the show-up profile's distribution, which we accomplish through a maximum likelihood estimation. The true likelihood function is intractable to work with, but we solve a much simpler approximate likelihood function. In Section 3, we provide theoretical justification for this method by proving that it results in consistent parameter estimates under mild assumptions.

Figure 3 presents one such show-up profile (for a single flight) obtained by assuming a truncated normal distribution and using maximum likelihood. This results in a mean earliness of 66 minutes with a standard deviation of 37 minutes. Compared with the industry show-up profile presented in Figure 1, forecasting arrivals using our method (which we refer to as a "structural" forecasting method) results in improved forecasting performance. Specifically, considering the root mean forecasting squared error, we find that the expected error across 15- minute intervals decreases by 13% using our method from a RMSE of 15.8 individuals to 13.7 individuals. We provide further details and analysis in Section 4.

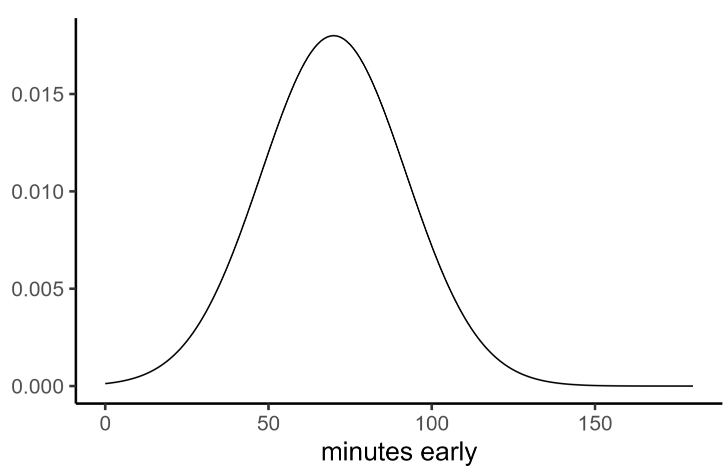

Figure 3 Example of a show-up profile trained using maximum likelihood estimation.

We further compare the performance of the structural forecasting method with black-box machine learning methods that use the time of day, day of week, multiple lagged passenger arrival counts, and multiple measures of the number of upcoming flights (for example, number of flights in the next 30 minutes, next 1 hour, etc.) and previous flights to predict future passenger arrivals. We present three main sets of performance results in Section 4. First, we show that the basic structural method performs similarly to machine learning methods, although in some cases slightly worse, and discuss reasons for this difference. Second, we demonstrate how the structural method can be improved by training multiple show-up profiles that differ based on flight characteristics such as departure time and destination. Finally, we show that improved forecasting performance can be obtained by combining machine learning methods, which can take into account many variables including recent lags, and structural methods, which take advantage of the special problem structure albeit in a static way. We accomplish this combination by using the estimate generated by the structural method as an additional feature variable for the machine learning methods. We find that we can sometimes improve machine learning methods by carefully reducing the set of feature variables to guide the methods to rely on relevant structure.

Our work demonstrates the estimation of arrival processes tied to show-up profiles and characterizes the consistency of the proposed estimators. The estimation approach can, in turn, be used to build structural forecasting methods that have performance and structural advantages over non-structural methods. Such a forecast is available far in advance of the focal period, and we find it to be robust across known schedule changes. Extensions of our approach allow the show-up profile to depend on flight features such as time of day or destination, and can therefore support empirical investigations of passenger behavior based on aggregate arrival counts, which we illustrate and discuss in Section 5. Finally, as noted earlier, the notion of a show-up profile has the distinct advantage of being familiar to and interpretable by practitioners.

The structure of the paper is as follows. We begin with a review of the relevant literature in Section 2. This section is followed by the formal problem formulation, development of our estimation method, and proof of the consistency of the estimation method in Section 3. Section 4 compares the performance of our method with that developed using the industry show-up profile estimate as well those under various machine learning methods. In Section 5, we present the results of our empirical analysis, and we provide our concluding remarks in Section 6. The appendix contains mathematical derivations, a discussion on consistency, and further robustness checks of numerical results.

2. Literature Review

Show-up profiles are known by various different names in the aviation industry, such as the passengers' arrival pattern (Gao et al. 2021), the lead time distribution (Kennon et al. 2013), and the earliness distribution (TransSolutions 2011), as well as a show-up profile (Isarsoft 2024). Show-up profiles are often used to predict congestion for resource planning as well as to better understand passenger behaviour. However, unlike in our work where we do not know the earliness of arriving passengers, the show-up profiles that are estimated in the literature assume access to exact measurements of passenger earliness, possibly by utilizing surveys or recording boarding pass information. Within the United States, boarding pass information is not recorded and stored (see Golden et al. 2007) and, more importantly for the purposes of our study, is not shared with the airport.

Airport congestion can be highly variable depending on the show-up profiles of the passengers that use the airport. Alodhaibi et al. (2019) looks at the impact that different arrival patterns have on congestion. Based on a model of the Brisbane International airport, the paper simulates the congestion at check-in counters, security, and immigration under various passenger show-up profiles. The authors find that the show-up profile's mean and maximum drastically affect the average and maximum amount of time spent waiting in each queue, especially when looking at times of high congestion. For example, the maximum amount of time spent waiting in security spans from 5.3 minutes to 23 minutes for the distributions they considered. This means that using the correct show-up profile is pivotal to having an accurate estimation of future arrivals.

Postorino et al. (2019) trains show-up profiles based on earliness measurements collected using Bar Coded Boarding Pass (BCBP) technologies to better understand passenger behavior. Similarly to our work, they train a parametric model of earliness and, as we also do in Section 5, they allow this model to differ based on various flight characteristics. The authors find that both the time of day and type of carrier (low-cost carrier or full carrier) affect the show-up profile, with low-cost carrier passengers arriving over a larger interval of time in the late afternoon period. Unlike our study, Postorino et al. have access to the exact earliness values of each individual passenger through the BCBP technology which greatly simplifies the estimation step. This technology is not used in the United States.

A common use of show-up profiles is in the estimation of passenger arrivals to determine optimal allocation of check-in counters, see for example Chun and Mak (1999), Al-Sultan (2018) and the survey in Lalita and Murthy (2022). Often, these show-up profiles are determined using passenger surveys, see Park and Ahn (2003), Ashford et al. (1976), Bruno et al. (2019), and Rauch and Kljajić (2006). As these studies assume access to instances of passenger earliness values, the estimation of a show-up profile is trivial. Unlike the approaches in this literature, our work does not have direct earliness measurements and thus cannot match arrivals to specific flights, which introduces unique challenges in estimation. However, these studies motivate the estimation of show-up profiles based on a number of factors, such as departure time, type of haulage (first-class and business class versus economy class passengers), and departure distance (domestic versus international flights) which we consider in Section 5.

Other studies, such as Ahyudanari and Vandebona (2005), use generic show-up profiles such as the distribution found in the Airport Development References Manual by the International Air Transport Association (IATA 1995) to estimate future arrivals. However, the IATA manual recommends that practitioners do their own survey to generate show-up profiles tailored to the airport being modeled. Our work deviates from this approach as we do not rely on surveys or previously existing profiles but instead estimate profiles directly from aggregated arrival data.

Forecasting future arrivals in airport settings goes beyond using show-up profiles. A stream of literature also considers using machine learning models and other models to forecast arrivals. However, as in the show-up profile models of the literature referenced above, these papers assume access to the historical earliness of each passenger. Sayın et al. (2023), for example, forecast arrivals using linear models and various types of recurrent neural nets that make predictions based on upcoming and past number of flight departures, distinguishing by type of aircraft (top large body, medium large body, etc). This study relies on access to the flight info of each passenger through boarding pass scans at the check-in counters, whereas we use aggregate arrival data without flight-specific identifiers.

Machine learning models are also used to estimate distributions of other elements of passenger flow such as connection times. Both Guo et al. (2020) and Guo et al. (2022) focus on modeling connecting passenger connection times with real-time data to improve multiple systems in the airport. These papers use regression trees to estimate the distribution of connection times, which is possible as they have access to information on every passengers and the flights (inbound and outbound) that they are associated with. Connection times, the region, time-of-day, day-of-week, and planned connection time are all considered as factors to forecast connection times.

Instead of using machine learning models or show-up profiles to forecast arrivals for check-in allocation, Parlar and Sharafali (2008) train a "show up process." They view passengers showing up for a flight as a death process. An arrival of a passenger from a fixed population, which encompasses all of the passengers that will be arriving, is viewed as an occurrence, i.e., a "death." For each passenger, the time until their arrival is assumed to follow an exponential distribution. The authors use a maximum likelihood technique to determine the rate of this death process. Our methodology similarly uses maximum likelihood estimation, but unlike Parlar and Sharafali (2008), we do not assume knowledge of specific passenger-to-flight mappings, instead we leverage aggregated arrival data only.

Forecasting arrivals is important in many settings other than airports, such as call centers, see Ibrahim and L'Ecuyer (2013), and hospitals and clinics, see Kim et al. (2015), among others. In particular, distributions of patient arrivals, similar to show-up profiles, are commonly used in healthcare settings to estimate how individuals will arrive for appointments or events. A common behavior to model is the incidence of no-shows or cancellations as in Liu et al. (2010). Dantas et al. (2018) provides a comprehensive literature review on the topic of no-shows. Unlike a healthcare environment, the uncertainty in passenger arrivals at airports is around the time of their arrival, rather than on the possibility of having no-shows. There are other papers that consider non-punctual customers, such as Jouini et al. (2022) and Koeleman and Koole (2012). While this literature provides valuable insights into predicting individual arrival behavior, it typically assumes knowledge of the specific event individuals are attending, which differs from our context in airport arrival forecasting, where we are not aware of which flight a passenger is arriving for.

3. Estimating a Show-up Profile

In this section, we develop a method to estimate show-up profiles of passenger arrivals. In Section 3.1, we introduce a model of passenger arrivals and provide a formal definition of a show-up profile. In Section 3.2, we construct a maximum likelihood estimate of the show-up profile, and in Section 3.3, we show that this estimate is consistent under mild assumptions.

3.1. A Show-up Profile of Passenger Arrivals

Suppose that we have passenger arrival data collected over M days, and on each day $m \in \{1, \dots, M\}$, we make observations over a bounded time range, $[0, T]$. Specifically, there are K scheduled flights on each day, where flight $k \in \{1, \dots, K\}$ is scheduled to depart at time $\tau_k \in [0, T]$. For expositional purposes, we make the assumption that flight schedules are identical across days. We relax this assumption in Sections 4 and 5, and we formally describe the relaxation in the appendix.

We assume that N_k, the number of passengers scheduled for flight k, is known and has the same value for all days. (In practice, airport planners can accurately estimate N_k by considering seat capacities on the corresponding flight and the expected load factor.) The total number of passengers to arrive on any given day is then equal to $\sum_{k=1}^{K} N_k$.

The arrival time of each passenger depends only on the passenger's scheduled flight time. Conditional on the passenger's scheduled flight time, the passenger's arrival time is independent of other flight times and the arrival times of other passengers. Furthermore, all passengers' earliness to their respective flights are stochastically identical. (In subsequent sections, we allow the show-up profiles to vary based on scheduled flight time and destination.) Specifically, each passenger arrives for their flight according to an identical show-up profile D_θ, where $D_\theta(t)$ represents the probability that a given passenger arrives more than t time units before their scheduled flight departure. The show-up profile D_θ follows a known parametric family characterized by a parameter θ which is contained in some subspace Θ of $\mathbb{R}^n$, $n \geq 1$. For a fixed θ, $D_\theta(t)$ is a survival function, or equivalently, $1 - D_\theta(t)$ is a cumulative distribution function (CDF). We suppose that passengers do not arrive at the TSA line after their respective scheduled flight time and arrive at most a time units before, where $a \in (0, \infty)$; for example, one can consider that passengers arrive no earlier than, say, $a = 360$ minutes prior to the scheduled departure of their flight. Therefore, $D_\theta(t) = 1$ for $t < 0$, and $D_\theta(t) = 0$ for $t \geq a$.

Our goal is to estimate θ, and thus obtain an estimate of the show-up profile D_θ, using arrival data that is aggregated over all flights. That is, we estimate θ even while we do not know which specific flight each arrival is associated with.

Each day is divided into I non-overlapping time intervals each of length r, for some fixed $r > 0$. These time intervals, which we call *sensor intervals*, correspond to data collection intervals of the people counters we installed at the airport. Sensor interval $i \in \{1, \ldots, I\}$ is defined as $(t_{i-1}, t_i]$, where $t_0 = 0$, $t_i = t_{i-1} + r$ for $i = 1, \ldots, I-1$, and $t_I = T \geq \tau_K$, i.e., the last flight departure occurs prior to the end of the last sensor interval. Let X_i^m denote the total number of arrivals (for any flight) during the sensor interval i, i.e., in the time interval $(t_{i-1}, t_i]$, on day m. Our objective is to estimate θ given x_i^m, the realization of random variables X_i^m for $i \in \{1, \ldots, I\}$ and $m \in \{1, \ldots, M\}$.

For expositional simplicity, we assume that flights depart on interval boundaries such that for each $k \in \{1, \ldots, K\}$, $\tau_k = t_i$ for some $i \in \{1, \ldots, I\}$. This assumption is also relaxed in our numerical experiments, and we again leave the modeling details to the appendix. We also assume that the first sensor interval occurs early enough that all arrivals are observed; namely, that the first flight of the day occurs at time a or later.

For any $i \in \{1, \ldots, I\}$, let

$$B_\theta(i) = D_\theta((i-1)r) - D_\theta(ir),$$

be the probability that a random passenger arrives in the ith sensor interval before their flight departure time; equivalently, $B_\theta(i)$ is the probability that the passenger arrives between $(i-1)r$ and ir minutes early relative to their flight departure time. We call the probability mass function $\{B_\theta(i),\, i = 1, \ldots, I\}$ the *binned distribution* or *binned show-up profile*. An example of a binned distribution for a passenger on flight 1 with a departure time of 7:30 is given in Figure 4. In the figure, the sensor interval size r is 15 minutes.

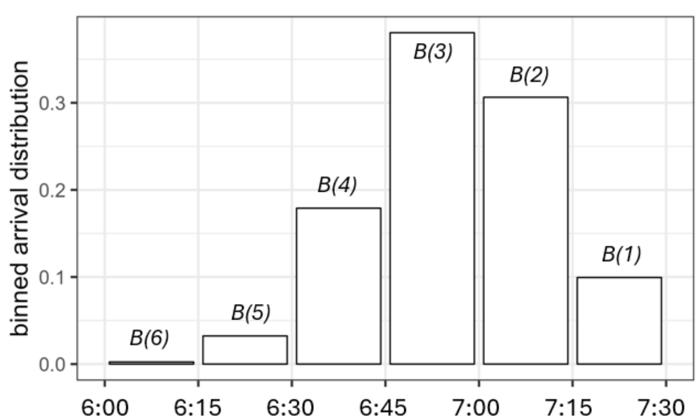

Figure 4 Example of a binned distribution with interval size of $r = 15$ minutes and flight departure time τ_1 of 7:30, where the x-axis represents the time of day. The binned distribution's dependence on θ has been suppressed for clarity of presentation.

3.2. Estimation of the Show-up Profile

Let int(k) denote the sensor interval in which the kth flight departs; i.e., int$(k) = j$ if $\tau_k = t_j$. In any day m, the probability that a passenger on the kth flight arrives in interval $i \leq$ int(k) is equal to $B_\theta(\text{int}(k) - i + 1)$, i.e., this is the probability that the passenger arrives in the $(\text{int}(k) - i + 1)$st interval before their flight departure. Also, the probability that a random passenger arriving on day m is on flight k is $N_k / \sum_{s=1}^{K} N_s$. Thus, by the law of total probability, the probability that a passenger on a fixed day m arrives in interval i, which we refer to as $P_\theta(i)$, is given by

$$P_\theta(i) = \frac{\sum_{k=1}^{K} B_\theta(\text{int}(k) - i + 1) N_k}{\sum_{k=1}^{K} N_k}.$$

For any fixed day m, we observe a collection of arrivals $X^m = \{X_i^m, \, i = 1, \ldots, I\}$. Note that the arrival observations $\{X^m, \, m = 1, \ldots, M\}$ are independent across days, as arrivals on each day are naturally independent. However, within a day, individual passenger arrival times are not independent. To see why there is a dependence across arrivals within a day, suppose we have observed N_K passengers arriving between times τ_{K-1} and τ_K. That is, we have seen as many passengers during the period of time between flight $K - 1$ (the next-to-last flight) and flight K (the last flight) as there are on flight K. Because we assume all passengers arrive before their scheduled flights, it follows that these passenger arrivals are the full set of arrivals for flight K. All other prior passenger arrivals must therefore belong to flights $1, \ldots, K - 1$, which results in a modified arrival time distribution. More generally, there is a complex dependence between passenger arrivals induced by the flight assignments. Explicitly expressing the true distribution of arrival vectors, much less calculating it, is quite complex. Hence, our goal of finding the parameter θ that drives these observations becomes intractable. For this reason, we now introduce a *surrogate problem* based on a mixture distribution.

In the surrogate problem, we do not enforce that each flight has a fixed number of passengers. Rather, we allow passengers to arrive for each flight i with probability $N_i / \sum_k N_k$, independently of each other. An alternative interpretation is that each passenger arrives according to a mixture distribution of the show-up profiles for all flights weighted by $N_i / \sum_k N_k$. Hence, as before, the probability a passenger arrives in interval i is equal to $P_\theta(i)$. In this surrogate problem, we let $\tilde{X}^m = \{\tilde{X}_i^m, \, i = 1, \ldots, I\}$ denote the counts of arrivals across all intervals on day m.

As each arrival can arrive independently in any of I potential sensor intervals, each with corresponding probability $P_\theta(i)$, $\tilde{X}^m = \{\tilde{X}_i^m, \, i = 1, \ldots, I\}$ follows a multinomial distribution with parameters $\sum_{k=1}^{K} N_k$, I, and $\{P_\theta(i), \, i = 1, \ldots, I\}$, corresponding to the number of trials, number of mutually exclusive events, and event probabilities, respectively. We can then express the likelihood of the show-up profile parameter ϑ as

$$L\left(\vartheta; \{\tilde{x}^m\}_{m=1}^{M}\right) = \prod_{m=1}^{M} \left(\frac{\left(\sum_{k=1}^{K} N_k\right)!}{\prod_{i=1}^{I} \tilde{x}_i^m!} \prod_{i=1}^{I} P_\vartheta(i)^{\tilde{x}_i^m} \right).$$

We can thus infer θ in the surrogate problem via maximum likelihood estimation as follows:

$$\tilde{\theta}_M = \arg\max_{\vartheta\in\Theta} \left\{ \sum_{m=1}^{M} \sum_{i=1}^{I} \tilde{x}_i^m \log P_\vartheta(i) \right\}. \quad (1)$$

Equation (1) is a conceptual construct because the real system does not produce observations $\tilde{x}_i^m$. Substituting the true observations x^m, we arrive at the following estimate:

$$\hat{\theta}_M = \arg\max_{\vartheta\in\Theta} \left\{ \sum_{m=1}^{M} \sum_{i=1}^{I} x_i^m \log P_\vartheta(i) \right\}. \quad (2)$$

Given the estimate $\hat{\theta}_M$ of θ, we also have an estimate of the binned distribution B_θ, given by $B_{\hat{\theta}_M}$, and an estimate of the show-up profile D_θ, given by $D_{\hat{\theta}_M}$.

3.3. Consistency of the Estimated Show-up Profile

In this section, we prove the consistency of estimator $\hat{\theta}_M$ in (2). The main result, presented in Theorem 1, guarantees that the parameter estimate $\hat{\theta}_M$ is weakly consistent under Assumption 1 below. We first present Proposition 1, which guarantees that $\tilde{\theta}_M$, the maximum likelihood estimate of the surrogate problem, converges almost surely to the true parameter θ.

ASSUMPTION 1.

1.1 *The true parameter θ lies within a compact subspace Θ of $\mathbb{R}^n$ where for each $\vartheta \in \Theta$, $B_\vartheta(i) = 0$ if and only if $B_\theta(i) = 0$.*

1.2 *The binned distribution $B_\vartheta(i)$ is continuous in ϑ for all i and $\vartheta \in \Theta$.*

1.3 *The parameter θ can be uniquely identified from the binned distribution B_θ.*

Assumptions 1.1 and 1.2 are mild regularity assumptions and are used to guarantee that the log likelihood is bounded and continuous. However, in general, the identifiability of the parameter θ from the binned distribution cannot be guaranteed under all scenarios. Certain parameter and interval configurations may lead to non-unique estimates based on the observed data alone. Even though we have constructed examples of non-identifiability (see Appendix B), the underlying scenarios are unrealistic and we have not faced the problem of non-identifiability in any of our numerical experiments.

One simple scenario in which the parameters are identifiable is when the show-up profile follows a discrete distribution which is equivalent to the binned distribution. Alternatively if the show-up profile follows a normal distribution, and the interval size is small enough such that $I \geq 3$, then the mean and standard deviation are identifiable. Likewise, if the show-up profile follows a triangular distribution with unknown mode, then the mode is identifiable as well. In Appendix B, we provide further discussion of the identifiability of the parameter θ from the binned distribution.

We now present our results on the consistency of estimators $\tilde{\theta}_M$ and $\hat{\theta}_M$ under Assumption 1. We show that the maximum likelihood estimator of the surrogate problem $\tilde{\theta}_M$ is strongly consistent in Proposition

1. Within the proof of this result, we show that Assumption 1.3, which states that the parameter θ can be uniquely identified from the binned distribution B_θ, guarantees that the parameter θ can be uniquely identified from P_θ as well. Proposition 1 follows from this result and the strong consistency of the maximum likelihood estimate of a multinomial distribution.

The main result, Theorem 1, is proven by first considering the consistency of the maximum value, rather than the consistency of the maximizer. This consistency is proven by considering two related stochastic processes and using the identifibility proven within Proposition 1 as well as Law of Large Numbers results. An Argmax Consistency Theorem is used to show the desired final result of the consistency of the maximizer.

PROPOSITION 1. *Under Assumption 1, the maximum likelihood estimate of the surrogate problem $\tilde{\theta}_M$ in (1) is strongly consistent.*

Proof. This proof is broken into two parts. First, we will show that we can derive a strongly consistent estimate $\hat{P}(i)$ of the probability of a random passenger arriving in any given sensor interval i, i.e. $\hat{P}_M$ is a strongly consistent estimator of P_θ. Second, we show that B_θ can be written as a linear function of P_θ; say $B_\theta(i) = g_i(P_\theta)$ for $i = 1, \ldots, I$. Thus, given a consistent estimate $\hat{P}_M$ of P_θ, the estimate $\hat{B}_M(i) = g_i(\hat{P}_M)$ will be a consistent estimator of $B_\theta(i)$. Finally, by Assumption 1.3, θ can be uniquely identified from B_θ, which implies that the $\tilde{\theta}_M$ for which $B_{\tilde{\theta}_M}(i) = \hat{B}_M(i) = g_i(\hat{P}_M)$ for all i is a consistent estimator for θ.

We first derive a consistent estimator for P_θ. Define $Y_i^M = \sum_{m=1}^M X_i^m$ to be the total number of passengers across all M days to arrive in sensor interval i. Recall that the observations for a single day m, $\tilde{X}^m$, follow a multinomial distribution with parameters P_θ, $\sum_{k=1}^K N_k$, and I, corresponding to the event probabilities, number of trials, and number of possible mutually exclusive events, respectively. Analogously, across multiple days, the sum of observations in the same period $Y^M = \{Y_i^M\}_{i=1,\ldots,I}$ follows a multinomial distribution with parameters P_θ, $M\sum_{k=1}^K N_k$, and I, corresponding to the event probabilities, number of trials, and number of possible mutually exclusive events, respectively.

As the maximum likelihood estimator of a multinomial distribution is strongly consistent, the estimate

$$\hat{P}_M = \arg\max_P \left\{ \sum_{i=1}^I \sum_{m=1}^M \tilde{x}_i^m \log P(i) : \sum_{i=1}^I P(i) = 1,\ P(i) > 0, \forall i \right\}$$

is a strongly consistent estimator of P_θ.

Now, given that the system of equations

$$P_\theta(i) = \frac{\sum_{k=1}^K B_\theta(\text{int}(k) - i + 1)N_j}{\sum_{k=1}^K N_k}, i = 1, \ldots I \tag{3}$$

has a unique solution $\{B_\theta(j)\}_{j=1}^I$, as $\hat{P}$ is a strongly consistent estimator of P_θ, we will have that the solution $\{\hat{B}(j)\}_{j=1}^I$ to

$$\hat{P}(i) = \frac{\sum_{k=1}^K \hat{B}(\text{int}(k) - i + 1)N_j}{\sum_{k=1}^K N_k}, i = 1, \ldots I$$

is a strongly consistent estimator of $\{B_\theta(j)\}_{j=1,\dots,I}$.

We will now write $B_\theta(i)$ as a linear function of P_θ for each $i = 1,\dots,I$. For notational simplicity, we will set the end time of the final sensor interval equal to the time of the final flight $T = t_I = \tau_K$. We will also assume that every period has exactly one departing flight. This can be done by allowing the total number of passengers on a flight to equal zero when there is no actual flight departing in that period, and writing the total number of passengers as a sum if there are multiple flights within that period. Thus, N_i corresponds to the number of passengers with a flight departing in sensor interval i. Note that this together simplifies the notation such that $I = K$ and causes the flight departure at the end of the day to be non-empty $N_I > 0$.

To show that equation (3) has a unique solution, we will solve backwards recursively, starting in the final sensor interval T, i.e., the interval of the final flight. We have the following probability of an arrival in the final interval I,

$$P_\theta(I) = \frac{B_\theta(1)N_I}{\sum_{k=1}^I N_k}$$

Thus, solving, we can obtain a solution for $B_\theta(1)$,

$$B_\theta(1) = \frac{P_\theta(I)\sum_{k=1}^I N_k}{N_I}.$$

Considering the prior interval $I-1$, we have that

$$P_\theta(I-1) = \frac{B_\theta(1)N_{I-1} + B_\theta(2)N_I}{\sum_{k=1}^I N_k}$$

which, solving for $B_\theta(2)$ results in

$$B_\theta(2) = \frac{P_\theta(I-1)\sum_{k=1}^I N_k - B_\theta(1)N_{I-1}}{N_I}.$$

More generally, for any $i = 0,\dots,I-1$,

$$P_\theta(I-i) = \frac{\sum_{j=0}^i B_\theta(i-j+1)N_{I-j}}{\sum_{k=1}^I N_k}$$

setting $N_i = 0$ for any $i \le 0$, resulting in the recursive solution

$$B_\theta(i+1) = \frac{P_\theta(I-i)\sum_{k=1}^I N_k - \sum_{j=1}^i B_\theta(i-j+1)N_{I-j}}{N_I}. \qquad \square$$

Consistency of $\tilde{\theta}_M$ ensures that as the number of observations M increases, $\tilde{\theta}_M$ converges to θ under the surrogate problem almost surely. We now turn our attention to the true observations and the estimator of primary interest, $\hat{\theta}_M$. We have the following result.

THEOREM 1. *Under Assumption 1, the estimate $\hat{\theta}_M \to \theta$ in probability as $M \to \infty$.*

Proof. Write for any possible parameter $\vartheta \in \Theta$ and vector $x \in \mathbb{R}^I$,

$$f(\vartheta, x) = \sum_{i=1}^{I} x_i \log P_\vartheta(i)$$

and

$$F(\vartheta) = \sum_{k=1}^{K} N_k \cdot \sum_{i=1}^{I} P_\theta(i) \log P_\vartheta(i).$$

As each day is independent, for each fixed i, X_i^m is independent across m. Moreover, as for each interval i and day m, $\mathbb{E}\left[X_i^m\right] = N_k \cdot P_\theta(i)$, and by the strong law of large numbers,

$$\lim_{M\to\infty} \frac{\sum_{m=1}^{M} f(\vartheta, X^m)}{M} = \mathbb{E}\left[f(\vartheta, X^1)\right] = F(\vartheta)$$

almost surely. For all M,

$$\begin{aligned}
&\left|\max_{\vartheta\in\Theta}\left\{\frac{\sum_{m=1}^{M} f(\vartheta, X^m)}{M}\right\} - \frac{\sum_{m=1}^{M} f(\theta, X^m)}{M}\right| \\
&= \left|\max_{\vartheta\in\Theta}\left\{\frac{\sum_{m=1}^{M} f(\vartheta, X^m)}{M} - F(\vartheta) + F(\vartheta)\right\} - \frac{\sum_{m=1}^{M} f(\theta, X^m)}{M}\right| \\
&\le \left|\max_{\vartheta\in\Theta}\left\{\frac{\sum_{m=1}^{M} f(\vartheta, X^m)}{M} - F(\vartheta)\right\} + \max_{\vartheta\in\Theta}\left\{F(\vartheta)\right\} - \frac{\sum_{m=1}^{M} f(\theta, X^m)}{M}\right| \\
&\overset{(a)}{=} \left|\max_{\vartheta\in\Theta}\left\{\frac{\sum_{m=1}^{M} f(\vartheta, X^m)}{M} - F(\vartheta)\right\} + F(\theta) - \frac{\sum_{m=1}^{M} f(\theta, X^m)}{M}\right| \\
&\le \left|\max_{\vartheta\in\Theta}\left\{\frac{\sum_{m=1}^{M} f(\vartheta, X^m)}{M} - F(\vartheta)\right\}\right| + \left|F(\theta) - \frac{\sum_{m=1}^{M} f(\theta, X^m)}{M}\right| \\
&\le \max_{\vartheta\in\Theta}\left|\frac{\sum_{m=1}^{M} f(\vartheta, X^m)}{M} - F(\vartheta)\right| + \left|F(\theta) - \frac{\sum_{m=1}^{M} f(\theta, X^m)}{M}\right|,
\end{aligned}$$

where (a) follows as θ can be uniquely identified from P_θ, as shown in the proof of Proposition 1, and as $\arg\max_{q\in\mathcal{P}} \sum_{i=1}^{I} p(i) \log q(i) = p$ for $p \in \mathcal{P} = \{p : \sum_{i=1}^{I} p_i = 1,\ p_i > 0, \forall i\}$.

Both of the terms above,

$$\max_{\vartheta\in\Theta}\left|\frac{\sum_{m=1}^{M} f(\vartheta, X^m)}{M} - F(\vartheta)\right| \quad \text{and} \quad \left|F(\theta) - \frac{\sum_{m=1}^{M} f(\theta, X^m)}{M}\right|,$$

converge in probability to zero. To see this, note that $\left|F(\theta) - \frac{\sum_{m=1}^{M} f(\theta, X^m)}{M}\right|$ converges almost surely, and hence also in probability, to zero as a result of the Strong Law of Large Numbers. Also, we can conclude that $\max_{\vartheta\in\Theta}\left|\frac{\sum_{m=1}^{M} f(\vartheta, X^m)}{M} - F(\vartheta)\right|$ converges in probability to zero by the Uniform Law of Large Numbers (ULLN). In particular, under Assumption 1, the ULLN as presented in Newey and McFadden (1994)'s Lemma 2.4[1] holds in our setting.

[1] Shi (2022) notes that that Lemma 2.4 is missing certain required measurability conditions. However, because of Assumption 1.2, our problem satisfies these additional criteria.

Finally, under Assumption 1, by the Argmax Continuous Mapping Theorem (see Theorem 3.2.2 in Vaart and Wellner 2023), and because

$$\left| \max_{\vartheta \in \Theta} \left\{ \frac{\sum_{m=1}^{M} f(\vartheta, X^m)}{M} \right\} - \frac{\sum_{m=1}^{M} f(\theta, X^m)}{M} \right|$$

converges in probability, we have that $\hat{\theta}_M \to \theta$ in probability, which concludes the proof. $\square$

Having established a consistent estimator of show-up profiles, we next turn our attention to forecasting using show-up profiles in Section 4 and to an investigation of how various factors correlate with passenger earliness in Section 5.

4. Forecasting

In this section we utilize the show-up profile introduced in Section 3 to forecast the future number of arrivals to the TSA security area given the schedule of flight departures. As this method will take advantage of the structure inherent to the arrival process, we will refer to this forecasting method as a *structural method*. To estimate show-up profiles, we utilize the methods described in Section 3.2, generalized to allow for differing flight schedules across days. We provide a detailed derivation of the generalized approach in Appendix A.

Given a parameter estimate $\hat{\theta}$ and an estimated binned show-up profile $B_{\hat{\theta}}$, we can forecast arrivals across a day by "stacking" the binned show-up profile in front of each scheduled flight departure time. That is, to forecast the number of passengers that will arrive in sensor interval j, we sum the probabilities that a passenger will arrive in sensor interval j, over all flights departing after the interval. To make this idea more precise, consider a future schedule of K^F flights departing at times $\{\tau_j^F\}_{j=1,\dots,K^F}$ with N_i^F passengers on the ith flight. As before, we assume that we observe all arrivals within a day (i.e., the first flight of the day occurs after time a and the final flight occurs before time T). Thus, for each sensor interval j, we obtain the forecasted number of arrivals as

$$\hat{A}(j) = \sum_{k=1}^{K^F} N_k^F B_{\hat{\theta}}(i^F(k) - j) \tag{4}$$

where the kth flight in the future flight schedule departs in time interval $i^F(k)$.

In practice, the number of passengers on each flight N_i^F needs to be estimated. As in Section 3, we estimate N_i^F by dividing the total number of passenger arrivals that day by the total number of flights. In practice, the airport has a fairly accurate estimate of the number of passengers per flight from information provided by the airlines or by simply multiplying the known capacity of each flight by an expected load factor (for example, 95%). Load factors themselves are relatively stable and predictable.

In this section, we investigate the predictive power of forecasts generated using the structural method. We seek to answer several questions through this analysis: (a) How does the predictive performance of the structural forecast compare with the performance of reasonable benchmark forecasting methods? (b) How can we extend the structural approach to improve its performance? (c) How can we combine the benefits of the structural forecasting approach with the benefits of other approaches?

4.1. Benchmark Forecasting Methods

As benchmarks for the structural forecasting method, we consider several time-series forecasting methods based on autoregressive and machine learning models. These include the auto-regressive model ARIMAX and several machine learning-inspired approaches including lasso and ridge regression, regression trees, bagging trees, and random forests. We train all benchmark models using a broad set of covariates including time of day and counts of future flights, past flights, and past passenger arrivals. Table 1 includes a full listing of covariates employed by our benchmark forecasting methods. Specifically, the variables **next_**$*$ count the number of flights scheduled to depart in an upcoming time period; e.g., next_60min counts the number of flights scheduled to depart in the next 60 minutes after the focal time period. The variables **prev_**$*$ count the number of flights scheduled to depart in a trailing period of time; e.g., prev_60min counts the number of flights scheduled to depart in the 60 minutes prior to the focal time period. The variables **lag_**$*$ count the number of passengers who arrived in a past 15-minute interval; e.g., lag_15min counts the number of arrivals in the 15-minute interval ending 15 minutes prior to the focal interval, while lag_1day_30min counts the number of passengers who arrived in the 15-minute interval that ended 1 day and 30 minutes prior to the focal time period.

We note that the structural forecast does not rely on lagged variables and therefore is not adaptive. That is, the structural point forecast is the same regardless of the forecast lead time. This is an advantage for implementation, as the structural approach can be calculated well in advance with a consistent set of covariates, but it is a potential drawback in terms of performance. We present one way to add adaptivity to the structural forecast in Section 4.5.

description	variable list
temporal	hour, weekday, month
number of upcoming flights	next_30min, next_60min, next_90min, next_120min, next_150min, next_180min
number of past flights	prev_30min, prev_60min, prev_120min
lagged arrival counts	lag_15min, lag_30min, lag_45min, lag_60min, . . . lag_120min, lag_1day_0min, lag_1day_15min, lag_1day_30min, . . . lag_1day_120min, lag_1week_0min, lag_1week_15min, lag_1week_30min, . . . lag_1week_120min, lag_2week_0min, lag_2week_15min, lag_2week_30min, . . . lag_2week_120min, lag_3week_0min, lag_3week_15min, lag_3week_30min, . . . lag_3week_120min, lag_4week_0min, lag_4week_15min, lag_4week_30min, . . . lag_4week_120min

Table 1 List of all variables available to the benchmark forecasting models.

For the benchmark methods, we consider forecast lead times ranging from 15-minutes (equivalent to one period in advance, given our sensors provide counts at 15-minute intervals) to four weeks (reflecting the planning horizon for certain staffing decisions). Longer lead times constrain the set of lagged covariates available to the benchmark forecasting methods. For example, for a 60-minute forecast lead time, the benchmark approaches only have access to lags less than 60 minutes. All lagged arrivals are available to the benchmark methods when considering a 15-minute lead time; no lagged arrivals are available to the benchmark methods when considering a 4-week lead time.

In the results that follow, we measure the accuracy of our point forecasts using root-mean squared error (RMSE) where, given true arrival observations $A(1), A(2), \ldots, A(T)$ and arrival forecast $\hat{A}(1), \hat{A}(2), \ldots, \hat{A}(T)$,

$$\text{RMSE} = \sum_{i=1}^{T} \left(A(i) - \hat{A}(i)\right)^2$$

for forecasting in intervals $1, 2, \ldots, T$. Our calculation omits the performance during the hours of 10:00pm and 2:00am when the TSA security area is idle. Results presented in this section are based on a training set ranging from January 1, 2022 until April 14, 2022 and a test set ranging from April 15, 2022 until May 9, 2022. We consider other training and test sets in Appendix C, where we find that our main qualitative findings are robust.

4.2. Configuring and Tuning

We seek to narrow the set of benchmark methods to simplify our later analysis. Table 2 reports the performance of several of the methods.

	forecast lead time							
method	15 minutes	1 hour	1 day	1 week	2 week	3 weeks	4 weeks	no lagged arrival data
lasso regression	12.03	12.77	12.87	13.12	13.92	14.43	14.80	15.68
ridge regression	12.21	12.78	12.85	13.18	13.98	14.49	14.83	15.61
bagging tree	12.34	12.37	12.45	12.52	12.81	13.15	13.20	13.62
random forest, $m = \sqrt{p}$	11.95	12.91	12.86	12.84	13.27	13.32	12.91	12.66
random forest, $m = p/3$	11.96	12.37	12.41	12.39	12.82	12.86	12.76	12.57
ARIMAX	12.35	17.00	21.04	23.61	23.61	23.61	23.61	23.61

Table 2 Performance of machine learning methods (RMSE) across various lead times.

The two penalized regression models, lasso and ridge regression, show similar performance across all lead times. In each model, the weight of the penalization is chosen using 5-fold cross validation. For clarity of presentation in our analysis going forward, we use ridge regression to represent these models.

In each random forest model, we build 100 trees and randomly sample m variables (out of the total number of variables p which depends on the lead time) as candidates at each split. We use both $m = \sqrt{p}$ and $m = p/3$, which are typical choices in the literature, see, for example, Gareth et al. (2013) and Breiman (2002). Bagging trees are random forests with m equal to the total number of variables p, or, equivalently, random forest models that allow all variables to be candidates at each split. The three tree-based methods exhibit similar performances to each other across various lead times, and we select the random forest with $m = p/3$ to highlight in our remaining results.

The random forest models perform similarly to the penalized regression models when the lead time is small. However, as the lead time increases, the random forest model outperforms the penalized regression models. To understand why, consider that the relationship between the recent lagged arrival variables and the upcoming number of arrivals is plausibly approximately linear while the relationship between the flight and temporal variables in relation to the upcoming number of arrivals is likely non-linear. It is reasonable that a weighted average of the recent lagged arrival data is a good predictor of upcoming number of arrivals, thus both sets of models do well when there is a short lead time. However, when recent lagged arrival data is not available, the models have to rely on flight and temporal data instead. As the random forest model allows for non-linear predictions, this gives an advantage to this model and results in better performance when the lead time is large.

Finally, we also consider an ARIMAX model – that is, an ARIMA model with regressors. For ARIMAX, rather than feeding the model all the lagged arrival counts displayed in Table 1, we give the model all past arrival counts and choose the number of past arrival counts to consider based on cross validation. This model compares similarly to others when predicting the number of arrivals in the next 15 minutes. However, When forecast with longer lead times, the ARIMAX model's performance degrades quickly. We note that as the lead time increases, we expect the ARIMAX model to behave similarly to an (unpenalized) linear regression model. It is dominated by the penalized regression models in this case. For this reason, we choose not to highlight ARIMAX as a benchmark in the results to follow.

For each of the structural forecasting methods, the show-up profile is assumed to follow a truncated normal distribution (truncated between 0 and 5 hours). We also considered other distributions and we show their performances in Table 3. As all performances were similar, we default to the truncated normal distribution.

distribution of show-up profile	RMSE
truncated normal distribution (min=0, max=5)	13.74
beta distribution (min=0, max=5)	13.83
triangular distribution	13.87

Table 3 RMSE when assuming show-up profiles that follow different parametric distributions.

4.3. Evaluating Predictive Performance

Our first insight is that the structural forecasting method performs similarly to the benchmark machine learning models, though slightly worse than the random forest model. Figure 5 gives the RMSE of the structural method, where the show-up profile is estimated through maximum likelihood, and of well-performing machine learning methods, ridge regression and random forest, across various lead times. The RMSE values furthest to the right on the x-axis correspond to forecasts computed with no lagged arrival counts. Each error can be viewed as the approximate expected error of forecasting arrivals in any 15 minute period between 2:00 am and 10:00 pm.

The structural method is invariant to lead time (hence, its RMSE appears as a straight horizontal line) as it estimates the number of arrivals based only on the show-up profile, an estimate of passenger counts

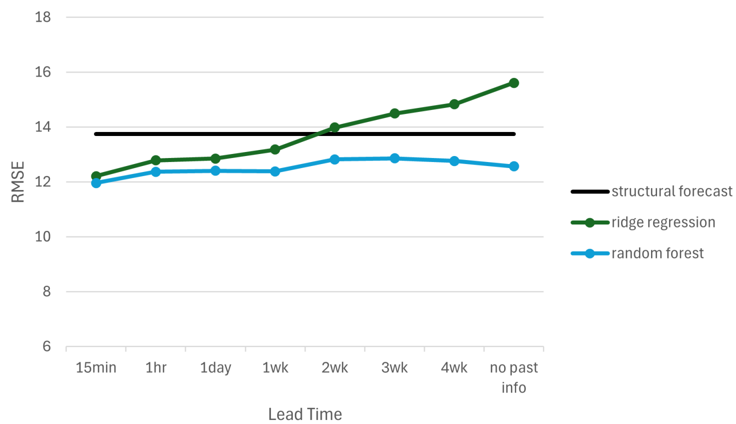

Figure 5 Performance comparison of truncated normal structural method, an in-practice method, ridge regression, and random forest. All RMSE are similar to one another with random forest performing the best.

on each flight, and the flight schedule within the test set. The structural method does not take into account lagged arrival counts and is thus invariant to the lead time.

The similarity in performance is surprising given how few variables the structural method uses versus the machine learning methods. While the structural method is only given the flight schedule and an estimated number of passengers on each flight, the machine learning methods are given the flight schedule, lagged arrival data, and temporal information. Even in the case of a long lead time or no lagged arrival data, the ability of the machine learning methods to adjust the forecast based on temporal information such as time-of-day and day-of-week should give the machine learning methods an advantage over the structural methods.

Looking at the performance of the random forest model in Figure 5, we notice that it levels off for long lead times and even slightly improves when there is no lagged data relative to 2- and 3-week forecast lead times. That is, the random forest model can actually benefit from having access to less data. Our interpretation of this observation is that for long lead-times the random forest approach is better off learning the relationship between passenger arrivals and flight departures than relying on outdated temporal passenger arrival patterns. Eliminating lagged passenger arrivals from the training set forces the random forest to learn the relationship between passenger arrivals and flight departures. This demonstrates that more data is not always better and that there can be advantages to using fewer variables which are known to be highly relevant according to a known structure.

4.4. Improving the Structural Methods

To motivate this subsection, consider training two different show-up profiles; one for passengers with flights departing before 9am, and another for passengers with flights departing after 9am. One way to do this is to allow the mean of each distribution to differ, which is displayed in Figure 6. Here, the standard deviation for each distribution is assumed to be equal and is estimated to be 37 minutes. The fitted distributions have means 54 and 75 minutes, for the pre-9am and post-9am flight departure times, respectively. That is, passengers arriving for early flights tend to arrive later, closer to their departure times, while passengers arriving for later flights tend to arrive farther in advance of their flights.

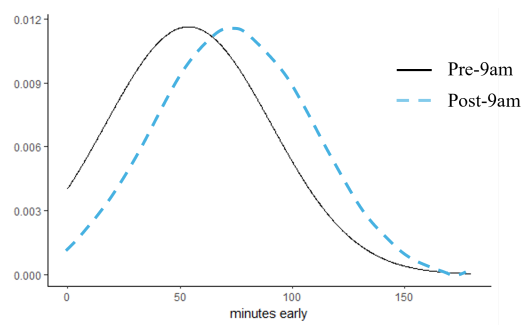

Figure 6 Two different show-up profiles fit for passengers arriving for flights scheduled to depart in the morning before 9:00am (solid curve) and flights scheduled after 9:00am (dashed curve). The pre-9:00am show-up profile shows passengers that arrive closer to their flight's departure time.

The difference between the two show-up profile means in Figure 6 is over 20 minutes, suggesting that we may achieve a substantially better model fit using two show-up profiles: one for pre-9am flights and another for post-9am flights. We find that this approach indeed yields an improvement in performance as detailed in Figure 7. We may try estimating different show-up profiles according to other characteristics. For example, Figure 7 shows that fitting different show-up profiles by both time-of-day (pre-9am versus post-9am) and destination yields a further improvement in out-of-sample forecast performance.

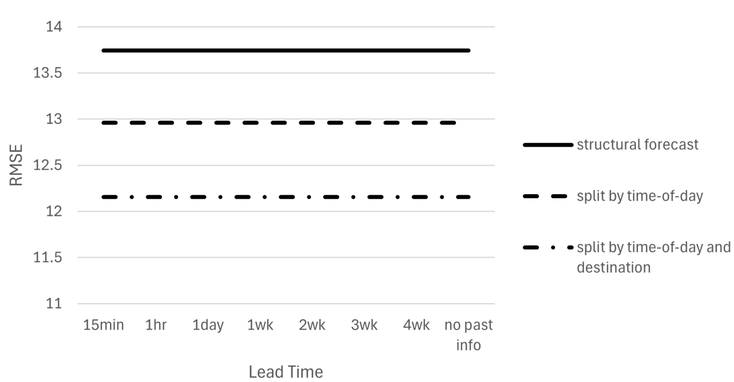

Figure 7 Performance comparison of the basic structural method, the structural method with pre-9am and post-9am profiles, and the structural method with profiles varying by time-of-day and destination.

Figure 6 and Figure 7 both motivate that the show-up profile of a passenger will depend on their flight's characteristics. We further explore this idea in Section 5.

4.5. Combining Structural Methods and ML for Best Performance

We next develop a method to take advantage of each technique's advantages. Namely, we would like to take advantage of the specific structure captured by the structural method and also the increased information and variable selection available to the machine learning methods. To do this, we first forecast using the structural method. Then, the forecasted estimates are fed to the machine learning models as an additional variable. We highlight the performance of this combined technique in Figure 8 alongside the structural forecast and random forest methods alone.

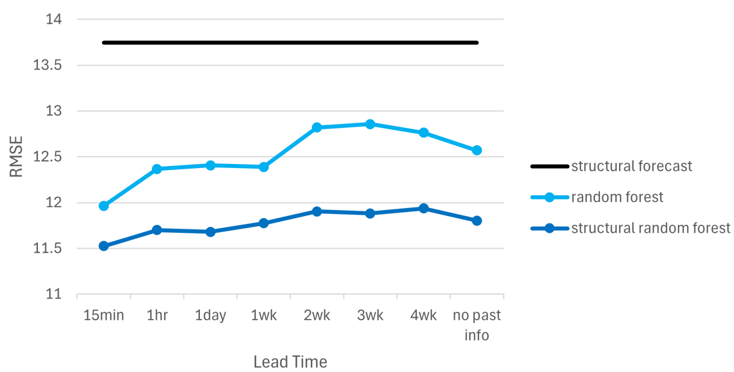

Figure 8 Performance comparison of the basic structural method, the random forest machine learning method, and a combined method.

By combining the structural and machine learning techniques, we allow the structural estimate to adapt to lagged arrival information and incorporate relevant temporal variables. Unlike the traditional machine learning techniques, however, incorporating the structural estimates relieves the machine learning method from discovering the relationship between flight patterns and arrivals; instead, this relationship is effectively summarized by a single informative covariate. As seen in Figure 8, this results in improved performance over each method used in isolation.

REMARK 1. Numerical analysis suggests that the parameter estimates of the show-up profile quickly converge. Even after only using one day of training data (in this example, April 14, 2022), we have the estimates of 65.7 minutes 32.4 minutes for the mean and standard deviation, respectively, compared to 66.0 minutes and 36.6 minutes when using 4 and a half months of training data (January 1, 2022 to April 14, 2022).

5. Analysis of Factors Impacting Show-Up Profiles

So far, we have utilized our capability to estimate show-up profiles exclusively to generate predictive models. However, the capability to estimate show-up profiles that vary with flight characteristics such as time-of-day and destination, demonstrated in Section 4.4, opens the possibility for using our approach to perform structural empirical analysis on the factors impacting passenger's arrival times before their flights. These insights, studied specific to an airport, can help managers with activities such as shaping passenger behavior (e.g., through targeted public service announcements encouraging passengers to arrive early) and scheduling flights (in which case understanding how adjacent flights might impact crowding at common service resources might be valuable). Our goal in this section is not to provide a comprehensive empirical analysis of show-up behaviors, but rather to demonstrate our methodology's usefulness for such an analysis. Our results in Section 4.4 suggest that show-up profiles may significantly differ for different passengers based on flight characteristics such as departure time and destination. Here, we more carefully study the relationships between these factors and passenger arrival behaviors.

As in Section 4, we assume the show-up profile of passengers follows a normal distribution truncated between 0 and 5 hours. We will assume all passengers share a show-up profile with identical standard deviation σ and a mean that can depend on characteristics of their respective flight. As before, we will estimate the show-up profile using the maximum likelihood estimate of Equation (2) using the same training set as in Section 4. We begin with an analysis of the effect of early morning departure times in Section 5.1 and continue with an extended analysis in Section 5.2 that considers multiple destination features.

5.1. The Effect of Early Morning Flights

Consider the effect that an early morning departure (departures at or before 9:00am) can have on the show-up profiles of passengers. We hypothesize that early morning departures result in passengers arriving significantly later, closer to their departure time. To motivate this hypothesis, we may consider that the passenger deciding when to arrive at the airport balances the cost of arriving early (which implies less time engaging in other activities) with the cost of missing their flight. For an early-morning flight, the cost of arriving early may be inflated because it requires waking up early and navigating morning traffic. In fact, differences in arrival behaviour to airports given early morning flights is considered in the literature, for examples, see Chun and Mak (1999), Postorino et al. (2019), and Park and Ahn (2003). This is also motivated by Figure 6 in which we see that our distributions appear dissimilar.

To test this hypothesis, we assume that for each passenger arriving for a respective flight with departure time t, the mean of their show up profile is equal to

$$\mu = \mu_{pre9am} + \mathbb{I}\{t > \text{9:00am}\} \cdot \alpha_{post9am}$$

where $\mathbb{I}\{s\}$ equals 1 if statement s is true and 0 otherwise.

Assuming each passenger's arrival is i.i.d., we estimate the parameters μ_{pre9am} and $\alpha_{post9am}$ along with the standard deviation σ using the maximum likelihood estimate of Equation (2). Table 4 gives the estimate for each parameter along with the standard error, estimated by using the Hessian of the likelihood function. We also give the t-statistic that corresponds to testing if the parameter $\alpha_{post9am}$ is zero; i.e., the mean of the

Show-up Profiles Split by Departure Time

	estimate	standard error	t-statistic
morning mean, μ_{pre9am}	53.72	0.24	
non-morning effect, $\alpha_{post9am}$	21.50***	0.31	69.12
standard deviation, σ	36.93	0.18	

Table 4 Parameter estimates and their respective standard errors for show-up profiles that differ for early-morning departures. Significance at the $\alpha = 0.10, 0.05,$ or 0.01 level is depicted using one, two, or three stars, respectively, on the corresponding estimate.

show-up profiles for early morning flights are the same as non-early morning flights. We also present the negative log likelihood value and the predictive performance on the test set presented in Section 4.

Note that this is equivalent to the show-up profiles trained in Section 4.4 with the same estimates and test performance. The non-early-morning effect $\alpha_{post9am}$ is significantly larger than zero, with early-morning passengers estimated to arrive on average 21.5 minutes later than passengers with flight departure times after 9:00am.

5.2. The Effect of Destination Characteristics

Motivated by the increased predictive performance seen in Section 4.4 when allowing the mean of the show-up profile to differ by destination, we consider a few possible explanations for what factors associated with destination may contribute to a significant change in average earliness.

We may hypothesize that leisure travelers, who tend to be less comfortable with navigating the airport, to arrive earlier than passengers traveling for business. This is supported by conversations with our airport contacts and appears in the literature as well (Ashford et al. 2011, Chun and Mak 1999, and Al-Sultan 2018). While we cannot determine the characteristics of passengers, we can classify destinations as either "business," "leisure," or "mixed." For the purposes of this analysis, the following classifications for the 14 different destinations seen in the training or test set were used; *leisure*: FLL, LAS, MCO, MIA, TPA; *mixed*: ATL, AUS, BNA, BWI, DEN, PHX; *business*: DAL, MDW, STL.

A second hypothesis is that the destination distance affects the show-up profiles of passengers. On one hand, passengers who are traveling a long distance may incur greater costs of missing their flights. In fact, Ashford et al. (2011) claims that "Generally the longer the distance traveled, the greater the time allowed by passengers prior to time of scheduled departure." On the other hand, passengers who are traveling a long distance are more likely to check bags, which may delay the time for them to reach the TSA security check-point. In the literature, distance is commonly accounted for by simply determining different show-up profiles for domestic versus international travel, see for example Ashford et al. (1976), Park and Ahn (2003), and Bruno et al. (2019). We look specifically at a terminal that only handles domestic travel to see if the distance to travel affects the show-up profile.

Third, we may consider whether a passenger's flight is to their final destination or whether they will be connecting to another flight. We may hypothesize that a passenger with a connecting flight may incur a higher cost of missing their initial flight. Again, this is something we cannot see in the data at the individual passenger level. However, we can try to control for this by categorizing destinations into those which commonly have connections and those who do not. During the dates of our training data, only one airline, Southwest, departed from the terminal studied. Southwest does not have official hubs. However, according to interviews (Russell 2020 and Ahlegren and Hardiman 2022) the following airports are commonly used for connections: ATL, BNA, BWI, DEN, MDW, PHX, and STL, which we label as *connecting* destinations. In the test data, there are also a few flights on the airline Spirit to the destinations MCO and MIA. We continue to consider these destinations, MCO and MIA, as non-connecting destinations for both airlines when evaluating test performance.

Taking into account all of the following variables along with the effect of early-morning departures, we model the average earliness as

$$\begin{aligned}\mu =& \mu_0 + \mathbb{I}\{t > \text{9:00am}\} \cdot \alpha_{post9am} + \mathbb{I}\{p = \text{mixed}\} \cdot \alpha_{mixed} + \mathbb{I}\{p = \text{business}\} \cdot \alpha_{business} \\ &+ \mathbb{I}\{c = \text{connecting}\} \cdot \alpha_{connecting} + d/100 \cdot \beta\end{aligned}$$

for a passenger with a corresponding flight with departure time t and destination with purpose p, connecting status c, and distance d.

Show-up Profiles Split by Destination Features

	estimate	standard error	t statistic
mean	63.91	0.54	
post-9am effect	22.86***	0.35	64.84
mixed destination effect	-6.47***	0.72	-8.99
work destination effect	-9.01***	0.74	-12.15
connecting effect	-5.66***	0.57	-9.88
increase per 100 miles	-0.035	0.0391	-0.898
standard deviation	38.73	0.19	

Table 5 Parameter estimates and their respective standard deviations for show-up profiles with mean varying based on various flight characteristics. Significance at the $\alpha = 0.10, 0.05,$ or 0.01 level is depicted using one, two, or three stars, respectively, on the corresponding estimate.

Note that, as before, our effect of early morning flights are significant. In fact, the effect is estimated to be similar to be with passengers who arrive for 9am or earlier flights arriving approximately 23 minutes later on average.

We also now consider the effect of destination by looking at (1) whether this destination is classified as leisure, work, or mixed; (2) whether this destination works as a hub for Southwest Airlines; and (3) the distance of the destination. We find that passengers traveling to leisure destinations arrive significantly sooner (i.e, further in advance of their flight) on average than those traveling to mixture or work destinations. This agrees with our hypothesis. On average, we see the leisure destination passengers arriving earlier by 6.47 compared to the mixture destination passengers and 9 minutes earlier than the work destination passengers.

We also find that those traveling to destinations that work as hubs for Southwest Airlines come significantly later on average (an average of 5.66 minutes later). This is in contrast to what we expected. One possible hypothesis for this is that passengers who are more likely to connect are also more likely to check their bags. As we are measuring time until reaching the TSA security area, checking a bag will delay a passenger's arrival to our sensors.

Finally, we find that destination distance, when looking at domestic travel only, does not appear to have a significant effect when controlling for other factors.

6. Conclusion and Future Research Directions

This paper demonstrates the estimation of show-up profiles from readily available data that does not match arriving passengers with departing flights. We present an estimation approach based on maximum likelihood that is consistent and efficient. Such an approach allows managers to estimate show-up profiles customized to particular settings, updated frequently, and accounting for relevant covariates. The resulting show-up profiles can be used to generate leadtime-independent structural forecasts whose quality rivals that of machine learning approaches based on a large set of time-series covariates in our airport setting. Furthermore, we show that by incorporating the structural forecasts into machine learning approaches, we obtain adaptive forecasts that significantly improve the machine learning approaches. It is clear that the structured information provided by the show-up profile model adds value even to sophisticated forecasting approaches. Show-up profiles bring the advantages that they are easy for airport managers to understand and to manipulate when conducting what-if analyses and when forecasting in new environments. We demonstrate that our estimation approach can also be used to test hypotheses on how passenger behavior depends on factors of interest.

As mentioned previously, while we tested our approaches in an airport setting where the notion of show-up profiles is known to be useful, we believe that show-up profiles and our methods have a broader set of potential applications in a more general set of "scheduled services" applications in which arrivals are driven by scheduled events. These include applications in other transportation settings, entertainment, and healthcare. A related problem is where customer arrivals are easily predicted and customers remain in an environment according to a "linger profile;" that is, a distribution of times in system. Variations of our methods may be used to estimate linger profiles and to forecast the census of customers in such a system over time.

Section 5 of our paper tests some basic insights into what factors correlate with passenger earliness. We believe our methods can be applied to perform more detailed empirical research on passenger earliness, for example to better understand heterogeneity among passengers and the mechanisms driving their arrival choices.

Acknowledgments

The authors are grateful for the support provided by the Raleigh-Durham Airport Authority and would like to thank in particular Bill Sandifer, Executive Vice-President and Chief Development Officer, and Delia Chi, Vice President of Planning and Sustainability, for their continued help throughout this research project. The authors also acknowledge helpful feedback from various seminar and conference participants.

References

Ahlegren, L. and Hardiman, J. (2022). Which airport does each major us carrier call home? *Simple Flying*. Available at: https://simpleflying.com/major-us-carrier-hubs/ (Accessed: September 26th, 2024).

Ahyudanari, E. and Vandebona, U. (2005). Simplified model for estimation of airport check-in facilities. *Journal of the Eastern Asia Society for Transportation Studies*, 6:724–735.

Al-Sultan, A. T. (2018). Simulation and optimization for modeling the passengers check-in system at airport terminal. *Review of Integrative Business and Economics Research*, 7(1):44.

Alodhaibi, S., Burdett, R. L., and Yarlagadda, P. K. (2019). Impact of passenger-arrival patterns in outbound processes of airports. *Procedia Manufacturing*, 30:323–330.

Ashford, N., Hawkins, N., O'Leary, M., Bennetts, D., and McGinity, P. (1976). Passenger behavior and design of airport terminals. *Transportation Research Record*, 588(3):18–26.

Ashford, N. J., Mumayiz, S., and Wright, P. H. (2011). *Airport engineering: planning, design, and development of 21st century airports*. John Wiley & Sons.

Bergin, M. (2024). Raleigh-Durham International Airport reports record passenger traffic in January 2024. *WRAL*. Available at: https://www.wral.com/story/raleigh-durham-international-airport-reports-record-passenger-traffic-in-january-2024/21285965 (Accessed: August 14th, 2024).

Breiman, L. (2002). Manual on setting up, using, and understanding random forests v3. 1. *Statistics Department University of California Berkeley, CA, USA*, 1(58):3–42.

Bruno, G., Diglio, A., Genovese, A., and Piccolo, C. (2019). A decision support system to improve performances of airport check-in services. *Soft Computing*, 23:2877–2886.

Cheng, L. (2014). *Modelling airport passenger group dynamics using an agent-based method*. PhD thesis, Queensland University of Technology.

Chun, H. W. and Mak, R. W. T. (1999). Intelligent resource simulation for an airport check-in counter allocation system. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 29(3):325–335.

Dantas, L. F., Fleck, J. L., Oliveira, F. L. C., and Hamacher, S. (2018). No-shows in appointment scheduling–a systematic literature review. *Health Policy*, 122(4):412–421.

Gao, Y. et al. (2021). Acrp problem statement: 454 describing, modelling and predicting passengers' arrival patterns. Technical report, AIRPORT COOPERATIVE RESEARCH PROGRAM.

Gareth, J., Daniela, W., Trevor, H., and Robert, T. (2013). *An introduction to statistical learning: with applications in R*. Spinger.

Golden, M. et al. (2007). Privacy impact assessment for the boarding pass scanning system. Technical report, U.S. Department of Homeland Security.

Guo, X., Grushka-Cockayne, Y., and De Reyck, B. (2020). London heathrow airport uses real-time analytics for improving operations. *INFORMS Journal on Applied Analytics*, 50(5):325–339.

Guo, X., Grushka-Cockayne, Y., and De Reyck, B. (2022). Forecasting airport transfer passenger flow using real-time data and machine learning. *Manufacturing & Service Operations Management*, 24(6):3193–3214.

IATA (1995). *Airport development reference manual*. International Air Transport Association.

Ibrahim, R. and L'Ecuyer, P. (2013). Forecasting call center arrivals: Fixed-effects, mixed-effects, and bivariate models. *Manufacturing & Service Operations Management*, 15(1):72–85.

Isarsoft (2024). What is a passenger show up profile? Available at: https://www.isarsoft.com/knowledge-hub/show-up-profile (Accessed: October 15th, 2024).

Josephs, L. (2024). Air travel demand is breaking records. airline profits aren't. *CNBC*. Available at: https://www.nbcnews.com/business/business-news/air-travel-demand-breaking-records-airline-profits-are-not-rcna160663 (Accessed: August 14th, 2024).

Jouini, O., Benjaafar, S., Lu, B., Li, S., and Legros, B. (2022). Appointment-driven queueing systems with non-punctual customers. *Queueing Systems*, 101(1):1–56.

Kennon, P., Hazel, R., Ford, E., and Hargrove, B. (2013). Acrp report 82: Preparing peak period and operational profiles—guidebook. Technical report, Transportation Research Board of the National Academies, Washington, D.C.

Kim, S.-H., Vel, P., Whitt, W., and Cha, W. C. (2015). Poisson and non-poisson properties in appointment-generated arrival processes: The case of an endocrinology clinic. *Operations Research Letters*, 43(3):247–253.

Koeleman, P. and Koole, G. (2012). Appointment scheduling using optimisation via simulation. In *Proceedings of the 2012 Winter Simulation Conference (WSC)*, pages 1–5. IEEE.

Lalita, T. and Murthy, G. (2022). The airport check-in counter allocation problem: A survey. *arXiv preprint arXiv:2208.13544*.

Liu, N., Ziya, S., and Kulkarni, V. G. (2010). Dynamic scheduling of outpatient appointments under patient no-shows and cancellations. *Manufacturing & Service Operations Management*, 12(2):347–364.

Newey, W. K. and McFadden, D. (1994). Large sample estimation and hypothesis testing. *Handbook of econometrics*, 4:2111–2245.

Park, Y. and Ahn, S. B. (2003). Optimal assignment for check-in counters based on passenger arrival behaviour at an airport. *Transportation Planning and Technology*, 26(5):397–416.

Parlar, M. and Sharafali, M. (2008). Dynamic allocation of airline check-in counters: a queueing optimization approach. *Management Science*, 54(8):1410–1424.

Postorino, M. N., Mantecchini, L., Malandri, C., and Paganelli, F. (2019). Airport passenger arrival process: Estimation of earliness arrival functions. *Transportation Research Procedia*, 37:338–345.

Rauch, R. and Kljajić, M. (2006). Discrete event passenger flow simulation model for an airport terminal capacity analysis. *Organizacija*.

Russell, E. (2020). Does southwest airlines have hubs? yes, but don't call them that. *The Points Guy*. Available at: https://thepointsguy.com/news/does-southwest-airlines-have-hubs/ (Accessed: September 26th, 2024).

Sayın, M. G., Aktaş, D. Y., Bolat, M., Çelenli, M. K., Dursun, B., Koç, G., and Üçkardeş, K. S. (2023). A study of predicting arrival patterns of airport passengers to the counters on the basis of international terminal. *Avrupa Bilim ve Teknoloji Dergisi*, (51):63–74.

Shi, X. (2022). Lecture 2: Some useful asymptotic theory. Available at: https://users.ssc.wisc.edu/ xshi/econ715/Lecture$_{2s}ome_{a}symptotic_{t}heorems.pdf(Accessed : October15th, 2024).$

TransSolutions (2011). Acrp report 55: Passenger level of service and spatial planning for airport terminals. Technical report, Transportation Research Board of the National Academies, Washington, D.C.

Vaart, A. v. d. and Wellner, J. A. (2023). Statistical applications. In *Weak Convergence and Empirical Processes: With Applications to Statistics*, pages 385–591. Springer.