MIT Open Access Articles

Predictive and Prescriptive Analytics Toward Passenger-Centric Ground Delay Programs

Predictive and Prescriptive Analytics toward Passenger-centric Ground Delay Programs

Alexandre Jacquillat

Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Massachusetts

Ground Delay Programs (GDP) comprise the main interventions to optimize flight operations in congested air traffic networks. The core GDP objective is to minimize flight delays, but this may result in optimal outcomes for passengers—especially with connecting itineraries. This paper proposes an original passenger-centric approach to GDP by leveraging data on passenger itineraries in flight networks. First, we identify analytical drivers of passenger-centric operations in transportation systems. Second, we develop an integer program that balances flight delays and passenger delays in large-scale GDP operations. A rolling procedure decomposes the problem while ensuring global feasibility, enabling the model's implementation in short computational times. Third, we propose statistical learning models to predict passenger itineraries and optimize GDP operations accordingly, enabling the model's implementation when passenger itineraries are unknown by air traffic managers. Computational results based on real-world data suggest that our modeling and computational framework can reduce passenger delays significantly at small increases in flight delay costs, and that these benefits are robust to imperfect knowledge of passenger itineraries. Results highlight two major levers of passenger-centric operations: (i) *delay allocation* (which flights to delay vs. prioritize), and (ii) *delay introduction* (whether to deliberately hold flights to avoid passenger misconnections).

Key words: Integer programming; Analytics; Passenger-centric operations; Air traffic management

1. Introduction

Most transportation networks involve two interconnected layers: vehicles and users. For instance, public transit systems operate subways and buses to transport riders; logistic systems operate ships, aircraft and trucks to transport packages; and air transportation systems operate aircraft to transport passengers. Extensive routing and flow management research has focused primarily on optimizing vehicle operations. But optimal vehicle operations might not lead to optimal outcomes from the end users' perspectives, especially when itineraries involve connections between multiple vehicles (e.g., multi-line transit itineraries, multi-modal logistics, and connecting flights).

This paper proposes an original passenger-centric approach to air traffic flow management, combining predictive and prescriptive analytics. The proposed approach leverages data on non-stop and connecting passenger itineraries to balance flight delays and passenger delays in large-scale networks of operations. In practice, however, passenger itineraries are not available to air traffic managers. We thus propose statistical learning models to predict passenger itineraries from historical data and optimize operations accordingly. Results suggest that, under perfect information,

1

passenger delays can be reduced very significantly (by 60–80% under good weather and by 20–60% under poor weather) through limited increases in flight delays (by 1–3%). Equally important, these benefits are very robust to imperfect information on passenger itineraries. Specifically, passenger-centric operations mitigate congestion costs by (i) selecting which flights to delay vs. prioritize, and (ii) deliberately adding departure delays to avoid passenger misconnections.

1.1. Background and Literature Review

Air Traffic Flow Management

Demand-capacity imbalances at busy airports can result in significant congestion, imposing costs in excess of $30 billion annually in the United States alone (Ball et al. 2010). The most obvious remedy—capacity expansion—is very expensive and not always feasible. A second option is to control peak-hour schedules through demand management (e.g., slot control, slot auctions, congestion pricing). These interventions can mitigate congestion (e.g., Swaroop et al. 2012, Vaze and Barnhart 2012, Jacquillat and Odoni 2015) but also impose economic costs due to output restrictions. Their implementation has thus been limited in the United States. A third option is to optimize flight operations through air traffic flow management (ATFM). This is the focus of this paper.

The core ATFM initiative is known as Ground Delay Programs (GDP). GDP optimize flights' departure and arrival times in capacitated networks. Their main objective is to absorb delays at departure airports rather than in the airspace, where they are most costly, challenging to operate and environmentally damaging. Odoni (1987) formalized the ground-holding problem. Richetta and Odoni (1993) and Terrab and Odoni (1993) optimized network-wide operations when a single airport is subject to GDP interventions. Vranas, Bertsimas, and Odoni (1994) tackled the multi-airport GDP problem. Ball et al. (2003), Mukherjee and Hansen (2007), Jones, Lovell, and Ball (2017) captured uncertainty on flight demand and airport capacities by means of stochastic optimization models. Bertsimas and Stock Patterson (1998), Bertsimas, Lulli, and Odoni (2011) developed ATFM models that optimize airport operations as well as en-route operations in capacitated air traffic control sectors. Thanks to advances in computing power, strong formulations and tailored algorithms, ATFM problems can now be solved for the full US and European networks (Lulli and Odoni 2007, Bertsimas, Lulli, and Odoni 2011, Balakrishnan and Chandran 2018).

In practice, GDP are now routinely implemented, with significant benefits for airlines, airports and passengers. A key success factor is the Collaborative Decision Making (CDM) paradigm, which fosters participation from the airlines through data sharing and communication (Ball et al. 2007, Vossen, Hoffman, and Mukherjee 2012). Two core CDM principles are to facilitate intra-airline and inter-airline slot exchanges (Wambsganss 1996, Vossen and Ball 2006, Pellegrini, Castelli, and Pesenti 2012) and to ensure inter-airline equity (Barnhart et al. 2012, Bertsimas and Gupta 2016).

All ATFM and CDM developments optimize flight delays and other flight-centric metrics such as equity (e.g., Barnhart et al. 2012), capacity utilization (e.g., Bertsimas and Gupta 2016), and service predictability (e.g., Ball et al. 2015). However, the ultimate costs of congestion do not depend solely on flight operations, but also on their effects on passenger itineraries. A first, and direct reason is that the same delays can induce higher costs if they are borne by flights carrying more passengers. To capture these effects, Vossen and Ball (2006) and Bertsimas, Farias, and Trichakis (2012) mentioned weighting each flight's delay by the number of passengers in it. But a second, more complex reason is that flight delays can disrupt connecting itineraries—through misconnections. Ball et al. (2010) and Barnhart, Fearing, and Vaze (2014) showed that passenger delays are mostly driven by a small fraction (2% to 5%) of itineraries being disrupted, especially when connecting airports are congested, when flights are full, and when re-accommodation opportunities are limited.

Evidence thus suggests that GDP outcomes could be improved by capturing passenger accommodations across flight networks. Yet, passenger-level information has not been incorporated into ATFM models. A likely reason is that passenger itineraries are known by airlines but not by air traffic managers. But Barnhart, Fearing, and Vaze (2014) built a database that estimates historical passenger flows in flight networks, based on a discrete-choice model. We leverage these data to develop a passenger-centric GDP approach that balances flight and passenger delays. We first assume perfect information, characterizing passenger-centric GDP operations in a hypothetical setting where passenger itineraries were shared by the airlines as part of the existing CDM environment. We then relax this assumption by using statistical learning to *predict* passenger itineraries and optimize GDP operations accordingly—thus characterizing passenger-centric GDP operations in the current environment where passenger itineraries are not available to air traffic managers.

Our results show that passenger delays can be reduced very significantly at limited costs in terms of flight delays. From a practical standpoint, this suggests opportunities to enhance GDP initiatives, even under imperfect information on passenger itineraries. Yet, the proposed approach does not intend to overhaul current decentralized practices, under which the airlines can re-optimize their operations as part of their disruption management processes. Instead, it aims to improve upstream ATFM decisions by making them more consistent with downstream passenger flows.

Data-driven decision-making

By combining predictive and prescriptive analytics, our approach falls into the broad umbrella of data-driven decision-making. This relates to many papers using a similar "Predict then Optimize" approach. Sample applications include supply chain management (Gallien et al. 2015), pricing (Ferreira, Lee, and Simchi-Levi 2016), assortment planning (Aouad, Farias, and Levi 2015), online advertising (Besbes, Gur, and Zeevi 2016), school bus routing (Bertsimas, Delarue, and

Martin 2019), disaster response (Dahan et al. 2020), smart cities (Liu, He, and Shen 2020), etc. Methodologically, Bertsimas and Kallus (2019) developed a stochastic optimization approach, by constructing data-driven scenarios that express each unknown parameter as a weighted average of training observations. In a similar spirit, Elmachtoub and Grigas (2017) proposed a "Smart Predict then Optimize" method where the learning model is trained—and evaluated—in view of the value of the optimization objective, as opposed to the prediction error.

This paper follows a similar philosophy—leveraging historical data to learn an uncertain component and to solve a downstream optimization problem accordingly. One particularity though is that, in our paper, uncertainty lies in the set of flight pairs with connecting passengers. From a predictive standpoint, this is treated as a categorical classification problem as opposed to a numerical regression problem. But from a prescriptive standpoint, the classifier is used to determine the *set* of active constraints and the resulting dynamics of the system—as opposed to a parameter.

1.2. Contributions and Outline

This paper develops an original passenger-centric approach to Ground Delay Programs. First, we identify analytical drivers of passenger-centric transportation operations. Second, we propose an integer programming model and a rolling horizon algorithm to optimize passenger-centric GDP operations in large-scale networks, assuming perfect information on passenger itineraries. Third, we predict passenger itineraries using statistical learning models, and optimize GDP operations accordingly. We show that the benefits of passenger-centric GDP operations are robust to data unavailability on passenger itineraries. Specifically, this paper makes the following contributions:

- *Deriving analytical insights on the trade-off between vehicle and passenger delays in transportation (Section 2).* We propose a Markov decision process that optimizes vehicle operations to balance passengers' wait times vs. misconnections. The optimal policy identifies two passenger-centric levers: (i) delay allocation (which vehicles to prioritize to minimize passenger delays), and (ii) delay introduction (whether to deliberately delay vehicles to avoid misconnections).

- *Formulating a bi-objective integer programming model that incorporates passenger accommodations into GDP optimization (Section 3).* The model optimizes flights' departure and arrival times in a network of capacitated airports, by trading off flight delay costs vs. passenger delays. It augments existing GDP models by capturing the impact of flight operations on passenger accommodations and delays, in the presence non-stop and connecting passenger itineraries.

- *Developing a dynamic rolling horizon algorithm to solve the passenger-centric GDP model efficiently (Section 4).* Consistently with the literature and practice, we optimize GDP operations with a rolling algorithm. At each hour, the model optimizes flight operations for a look-ahead window, while tracking passenger accommodations across the day. We add constraints to

ensure consistency in passenger accommodations and global feasibility. The algorithm yields high-quality solutions in short runtimes consistent with practical requirements.

- *Characterizing the benefits of passenger-centric GDP using real-world data (Section 5).* We implement our model at the 30 busiest US airports. Results show that, as compared to a flight-centric baseline, passenger delays can be reduced very significantly (by 60–80% under good weather and by 20–60% under poor weather) through comparatively small increases in flight delay costs (by 1–3%). Results also provide interpretable guidelines to capture passenger-level objectives: flights with many non-stop passengers, many outgoing connections, and few incoming connections should be prioritized, while others should be de-prioritized or delayed.

- *Developing a predictive model of passenger itineraries, and showing the robustness of our passenger-centric approach to data unavailability (Section 6).* Using historical data, we use predictive analytics to infer the number of non-stop passengers on each flight and the flight pairs with connecting passengers. The models achieve strong out-of-sample performance, with an R^2 of 0.945 and an area under the curve of 95%, respectively. To our knowledge, this is the first application of statistical learning methods to predict passenger itineraries in flight networks. Results suggest that the benefits of passenger-centric GDP operations are extremely robust to imperfect information on passenger itineraries: the GDP model based on predicted itineraries achieves most of the benefits of the one based on perfect information.

## 2. Analytical Model of Passenger-centric Operations

We propose a continuous-time Markov decision process, capturing a general setting where vehicles are operated at a facility to transport passengers (or packages). We characterize the optimal policy to identify the main drivers of passenger-centric operations. All proofs are reported in Appendix A.

2.1. Model Development

Consider a facility where *arriving vehicles* are incoming and *departing vehicles* need to be operated. We consider (i) *departing passengers* (in a departing vehicle), (ii) *arriving passengers* (in an arriving vehicle), and (iii) *connecting passengers* (transferring from an arriving vehicle to a departing vehicle). We optimize which vehicle to operate, if any, to minimize the delay borne by all—departing and connecting—passengers. All costs are discounted at rate $\beta > 0$.

Notations are shown in Figure 1. Let n and m denote the number of vehicles yet to depart and arrive, respectively. We index them by $i, k = 1, \cdots, n$ and $j, l = 1, \cdots, m$. Let γ_{ij} be the number of passengers connecting from vehicle i to vehicle j. We assume that the arrival of each incoming vehicle i follows an exponential distribution with rate λ_i. Similarly, the service time to operate each departing vehicle j follows an exponential distribution with rate μ. Each passenger misconnecting on departing vehicle j bears a cost c_j (in time units). Practically, λ_i, μ and c_j reflect the scheduled

arrival of vehicle i, the facility's capacity, and the inverse frequency of vehicles substituting vehicle j, respectively. We assume that $\lambda_i < \mu$ for all i (it is faster to operate a departure than to wait for the next arrival) and $c_j > \frac{1}{\mu}$ for all j (it is cheaper to operate a vehicle than to miss a connection).
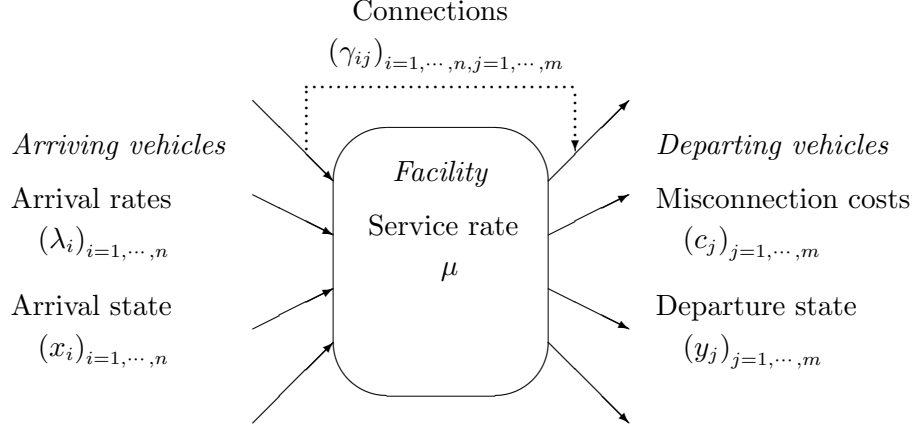
Figure 1   Setting and notations for the analytical model.

The state variable, denoted by (x, y), characterizes which vehicles have arrived and departed, and tracks the number of passengers in departing vehicles. It is defined as follows:

$$x_i = \begin{cases} 1 & \text{if vehicle } i \text{ is yet to arrive} \\ 0 & \text{otherwise} \end{cases}$$

$$y_j = \begin{cases} \text{number of passengers in vehicle } j \text{ if vehicle } j \text{ is yet to depart} \\ T \text{ otherwise (terminal state)} \end{cases}$$

Let $\mathcal{A}(x)$ and $\mathcal{B}(y)$ denote the set of vehicles yet to arrive and depart, respectively:

$$\mathcal{A}(x) = \{i = 1, \cdots, n \mid x_i = 1\} \quad \text{and} \quad \mathcal{B}(y) = \{j = 1, \cdots, m \mid y_j \neq T\}$$

For each state (x, y), we denote by $f_i(x, y)$ and $g_j(x, y)$ the state following the arrival of vehicle $i \in \mathcal{A}(x)$ and the departure of vehicle $j \in \mathcal{B}(y)$, respectively. Upon arrival of vehicle i, we track all passengers that connect from vehicle i to all the departing vehicles $j \in \mathcal{B}(y)$. Upon the departure of vehicle j, the arrival vector is unchanged. We thus have:

$$f_i(x, y) = (\widehat{x}, \widehat{y}) \quad \text{with:} \quad \begin{cases} \widehat{x}_i = 0 \\ \widehat{x}_k = x_k, \quad \forall k \neq i \end{cases} \quad \text{and:} \quad \begin{cases} \widehat{y}_j = y_j + \gamma_{ij}, & \forall j \in \mathcal{B}(y) \\ \widehat{y}_j = T, & \forall j \notin \mathcal{B}(y) \end{cases}$$

$$g_j(x, y) = (x, \bar{y}) \quad \text{with:} \quad \begin{cases} \bar{y}_j = T \\ \bar{y}_l = y_l, \quad \forall l \neq j \end{cases}$$

We denote by $\pi^*(x, y)$ the optimal policy and by $J^*(x, y)$ the optimal cost-to-go function. Each time a vehicle arrives to or departs from the facility, the decisions involve operating a departing vehicle $j \in \mathcal{B}(y)$ (referred to as $\pi^*(x, y) = j$) or operating no vehicle (referred to as $\pi^*(x, y) = 0$).

The objective of the model is to minimize total passenger delays. Lemma 1 elicits the discrete Bellman equation for this problem. It includes: (i) the continuous delay cost borne by all passengers

in $\mathcal{B}(y)$ waiting to depart, (ii) the costs of misconnections upon vehicle arrivals, (iii) the cost-to-go upon the next arrival, and (iv) the cost-to-go reflecting the decision of operating no vehicle (and staying in (x, y) at rate μ) or operating vehicle $l \in \mathcal{B}(y)$ (and transitioning to $g_l(x, y)$ at rate μ).

LEMMA 1. *The Bellman equation of the model is given by:*

$$J^*(x, y) = \frac{1}{\beta + \mu + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left[\sum_{j \in \mathcal{B}(y)} y_j + \sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j \notin \mathcal{B}(y)} \gamma_{ij} c_j \right) + \sum_{i \in \mathcal{A}(x)} \lambda_i J^* \left(f_i(x, y) \right) \right.$$
$$\left. + \mu \min \left\{ J^*(x, y), \min_{l \in \mathcal{B}(y)} \left(J^* \left(g_l(x, y) \right) \right) \right\} \right] \tag{1}$$

We first derive the optimal policy in small-scale settings (with one arriving vehicle and up to two departing vehicles) to identify the main levers employed in passenger-centric operations. Results also show that the problem does not admit a simple index policy. We thus focus next on the high-capacity regime where $\mu \to \infty$, i.e., when the time to operate each vehicle is infinitesimally small compared to the waiting times to avoid misconnections and to the costs of misconnections.

2.2. Optimal Policy in Small-scale Settings

Proposition 1 elicits the optimal policy when $n = m = 1$. When the incoming vehicle has arrived, the departing vehicle should always be operated—there is no more value in waiting. Otherwise, the departing vehicle is operated if and only if $y_1 \geq \frac{\lambda_1 \gamma_{11}(c_1(\beta + \mu) - 1)}{\beta + \mu + \lambda_1}$. That is, it is more beneficial to operate the departing vehicle as (i) it carries more departing passengers y_1, (ii) it is waiting for fewer incoming connections γ_{11}, (iii) the costs of waiting are higher (i.e., the arrival rate λ_1 is smaller and/or the capacity μ is smaller), and (iv) the misconnection cost c_1 is lower. Otherwise, it is more beneficial to hold the departing vehicle until the arrival of the incoming vehicle.

PROPOSITION 1. *If $n = m = 1$, then for any $y \geq 0$, the optimal policy is given by:*

$$\pi^*(0, y) = 1$$
$$\pi^*(1, y) = \begin{cases} 1 & \text{if } y_1 \geq \frac{\lambda_1 \gamma_{11}(c_1(\beta + \mu) - 1)}{\beta + \mu + \lambda_1} \\ 0 & \text{otherwise} \end{cases}$$

Proposition 2 turns to the case where $n = 1$ and $m = 2$. Once the incoming vehicle has arrived, the optimal policy simply prioritizes the vehicle with most passengers. Prior to the incoming vehicle's arrival, the optimal policy prioritizes the vehicle j with the highest value of $y_j - \frac{\lambda_1 \gamma_{1j}(c_j(\beta + \mu) - 1)}{\beta + \mu + \lambda_1}$. As in Proposition 1, this reflects the difference between delay vs. misconnection costs. But there are two novelties. First, a vehicle might depart even if the cost of misconnections exceeds the delay cost. Consider the case where $0 > y_1 - \frac{\lambda_1 \gamma_{11}(c_1(\beta + \mu) - 1)}{\beta + \mu + \lambda_1} \geq y_2 - \frac{\lambda_1 \gamma_{12}(c_2(\beta + \mu) - 1)}{\beta + \mu + \lambda_1}$ and $y_1 - \frac{\lambda_1 \gamma_{11}(c_1(\beta + \mu) - 1)}{\beta + \mu + \lambda_1} + \frac{\lambda_1 \mu \min(y_1 + \gamma_{11}, y_2 + \gamma_{12})}{(\beta + \mu)(\beta + \mu + \lambda_1)} \geq 0$. If vehicle 1 was the only vehicle, we would wait for the incoming arrival

(Proposition 1). But with two departing vehicles, we operate vehicle 1 (Proposition 2): the facility's limited capacity results in higher delay costs with multiple departing vehicles, hence in a more aggressive policy. Second, the priority order between the two departing vehicles may change upon the arrival of the incoming vehicle. This occurs, for instance, if $y_1 - \frac{\lambda_1 \gamma_{11}(c_1(\beta+\mu)-1)}{\beta+\mu+\lambda_1} > y_2 - \frac{\lambda_1 \gamma_{12}(c_2(\beta+\mu)-1)}{\beta+\mu+\lambda_1}$ and $y_1 + \gamma_{11} < y_2 + \gamma_{12}$. In this case, the facility prioritizes vehicle 1 before the incoming vehicle's arrival, but vehicle 2 after the incoming vehicle arrives.

PROPOSITION 2. *If $n=1$ and $m=2$, then for any $y_1, y_2 \geq 0$, the optimal policy is given by:*

$$\pi^*(0, y_1+\gamma_1, y_2+\gamma_2) = \begin{cases} 1 & if\ y_2+\gamma_{12} \leq y_1+\gamma_{11} \\ 2 & if\ y_1+\gamma_{11} \leq y_2+\gamma_{12} \end{cases}$$

$$\pi^*(1, y_1, y_2) = \begin{cases} 1 & if\ y_1 - \frac{\lambda_1 \gamma_{11}(c_1(\beta+\mu)-1)}{\beta+\mu+\lambda_1} \geq y_2 - \frac{\lambda_1 \gamma_{12}(c_2(\beta+\mu)-1)}{\beta+\mu+\lambda_1} \\ & and\ y_1 - \frac{\lambda_1 \gamma_{11}(c_1(\beta+\mu)-1)}{\beta+\mu+\lambda_1} + \frac{\lambda_1 \mu \min(y_1+\gamma_{11}, y_2+\gamma_{12})}{(\beta+\mu)(\beta+\mu+\lambda_1)} \geq 0 \\ 2 & if\ y_2 - \frac{\lambda_1 \gamma_{12}(c_2(\beta+\mu)-1)}{\beta+\mu+\lambda_1} \geq y_1 - \frac{\lambda_1 \gamma_{11}(c_1(\beta+\mu)-1)}{\beta+\mu+\lambda_1} \\ & and\ y_2 - \frac{\lambda_1 \gamma_{12}(c_2(\beta+\mu)-1)}{\beta+\mu+\lambda_1} + \frac{\lambda_1 \mu \min(y_1+\gamma_{11}, y_2+\gamma_{12})}{(\beta+\mu)(\beta+\mu+\lambda_1)} \geq 0 \\ 0 & if\ \max\left\{ y_1 - \frac{\lambda_1 \gamma_{11}(c_1(\beta+\mu)-1)}{\beta+\mu+\lambda_1}, y_2 - \frac{\lambda_1 \gamma_{12}(c_2(\beta+\mu)-1)}{\beta+\mu+\lambda_1} \right\} \\ & \quad + \frac{\lambda_1 \mu \min(y_1+\gamma_{11}, y_2+\gamma_{12})}{(\beta+\mu)(\beta+\mu+\lambda_1)} < 0 \end{cases}$$

These results outline two main strategies employed to minimize passenger costs. First, *delay allocation* prioritizes departing vehicles to minimize the impact of delays on passenger wait times. It also involves delay re-allocation, i.e., updates in the priority order upon the arrival of incoming vehicles based on the number of realized connections. Second, *delay introduction* deliberately adds departure holds to avoid misconnections, when resulting benefits outweigh added delay costs.

2.3.   Optimal Policy with $\mu \to \infty$

Practically, the high-capacity regime is motivated by the small time required to operate departures, as compared to waiting for arrivals or re-accommodating misconnecting passengers. Technically, it ensures tractability by avoiding operating vehicles with a "deficit" of departing vs. connecting passengers (as in Proposition 2)—instead, the optimal policy follows a simpler index rule.

By taking the limit in Equation (1) as $\mu \to \infty$, we obtain Equation (2). The first term corresponds to the instantaneous transition from state (x, y) to state $g_l(x, y)$ if vehicle $l \in \mathcal{B}(y)$ is operated. The second term captures the expected cost if no vehicle is operated, including passenger delays, misconnection costs, and the cost-to-go from vehicle i's arrival onward.

$$J^*(x, y) = \min \left\{ \min_{l \in \mathcal{B}(y)} J^*\left(g_l(x, y)\right), \right.$$
$$\left. \frac{1}{\beta + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left[\sum_{j \in \mathcal{B}(y)} y_j + \sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j \notin \mathcal{B}(y)} \gamma_{ij} c_j \right) + \sum_{i \in \mathcal{A}(x)} \lambda_i J^*\left(f_i(x, y)\right) \right] \right\} \quad (2)$$

In the high-capacity regime, the decision involves "which vehicles to operate vs. hold" rather than "which vehicles to prioritize". Indeed, operating vehicles j and l is equivalent to operating

vehicles l and j. We thus denote by $\Pi^*(x, y)$ the *set* of optimal decisions in state (x, y). Consistently, we focus here on the *delay introduction* insights rather than on the *delay allocation* insights.

The optimal policy is elicited in Proposition 3. The main result is that departing vehicle $j \in \mathcal{B}(y)$ should be operated if and only if $y_j \geq \sum_{i \in \mathcal{A}(x)} \lambda_i \gamma_{ij} c_j$, i.e., if the cost borne by the passengers waiting in vehicle j exceeds the cost of misconnections (across all arriving vehicles). The second part of the proposition simply asserts that the system instantaneously transitions from state (x, y) to a state $\delta(x, y)$ where all the vehicles $j \in \mathcal{B}(y)$ satisfying $y_j \geq \sum_{i \in \mathcal{A}(x)} \lambda_i \gamma_{ij} c_j$ have departed. The proof is more technical, and proceeds by double induction over $|\mathcal{A}(x)|$ and $|\mathcal{B}(y)|$.

PROPOSITION 3. *We denote by* $\mathcal{D}(x, y) = \left\{ j \in \mathcal{B}(y) \mid y_j < \sum_{i \in \mathcal{A}(x)} \lambda_i \gamma_{ij} c_j \right\}$, *and by* $\delta(x, y)$ *the m-vector defined by* $[\delta(x, y)]_j = y_j$ *if* $j \in \mathcal{D}(x, y)$ *and* $[\delta(x, y)]_j = T$ *if* $j \notin \mathcal{D}(x, y)$. *We have:*

$$\Pi^*(x, y) = \mathcal{B}(y) \setminus \mathcal{D}(x, y) \qquad i.e., \quad j \in \Pi^*(x, y) \iff y_j \geq \sum_{i \in \mathcal{A}(x)} \lambda_i \gamma_{ij} c_j$$

$$J^*(x, y) = J^* (\delta(x, y)) = \frac{1}{\beta + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left\{ \sum_{j \in \mathcal{D}(x,y)} y_j + \sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j \notin \mathcal{D}(x,y)} \gamma_{ij} c_j \right) + \sum_{i \in \mathcal{A}(x)} \lambda_i J^* (f_i(x, \delta(x, y))) \right\}$$

This result shows that the benefits of operating vehicle $j \in \mathcal{B}(y)$ increase with y_j, but decrease with λ_i, γ_{ij} and c_j. This extends the insights from Section 2.2 to a setting with several vehicles. Moreover, the benefits of operating any vehicle $j \in \mathcal{B}(y)$ increase whenever an incoming vehicle $i \in \mathcal{A}(x)$ arrives—the cost of misconnections decreases by $\lambda_i \gamma_{ij} c_j$. The optimal policy is thus to wait for incoming vehicles just until the number of passengers ready to depart in vehicle j exceeds $\sum_{i \in \mathcal{A}(x)} \lambda_i \gamma_{ij} c_j$, and operate vehicle j immediately thereafter.

2.4. Synthesis of Insights

Passenger-centric operations involve two strategies: (i) *delay allocation* (i.e., which departures to prioritize), and *delay re-allocation* (i.e., whether to update priorities upon vehicle arrivals), and (ii) *delay introduction* (i.e., whether to deliberately hold departing vehicles to avoid misconnections). Delay allocation and re-allocation reduce passenger delays without adding vehicle delay, whereas delay introduction trades off vehicle vs. passenger delays. Results also identified the main drivers underlying passenger-centric operations: any departure is more beneficial with more departing passengers, fewer incoming connections, closer incoming vehicle arrivals, and less costly misconnections. As we shall discuss, these insights will be applicable in the more complex GDP setting.

3. Integer Programming Model of Passenger-centric GDP

We now develop an integer programming model, referred to as (GDP-PAX), to optimize passenger-centric GDP operations. Unlike in our analytical setting, the decision to operate vs. delay a departing flight is now complicated by combinatorial complexities arising in air traffic networks. Consider

for instance a flight with many incoming connections. From Section 2, one may wish to hold this flight to avoid misconnections if inbound flights are delayed. But this decision now depends on ripple effects on airport, aircraft and passenger networks. For the airport, flight schedules may be such that operations will be capacity-constrained shortly thereafter, so any departure hold might create further delays. For an aircraft, a departure hold might induce further delays at the destination if the turnaround time is insufficient. For passengers, the departure hold may enable more connections at the origin airport, but may also create new misconnections at the destination airport if the flight was booked as the second leg by some passengers, but as a first leg by others. The (GDP-PAX) model captures these interdependencies across networks of operations.

In this paper, we focus on the deterministic multi-airport ground-holding problem. This restriction enables us to capture the trade-off between flight and passenger delays, while retaining tractability. Our passenger-centric approach can be integrated in future research into ATFM models that capture en-route capacities (e.g. Bertsimas and Stock Patterson 1998, Bertsimas, Lulli, and Odoni 2011) and operating uncertainty (e.g. Ball et al. 2003, Balakrishnan and Chandran 2018).

3.1.   Presentation

The (GDP-PAX) model takes as inputs flight schedules, aircraft connections, airport capacities, and passenger itineraries. Flight schedules and aircraft connections are available from the published schedule or updated flight plans. Airport capacities can be estimated from historical data. We first assume that passenger itineraries are known by air traffic managers—for instance, they can be provided by the airlines as part of the Collaborative Decision Making (CDM) program. We relax this assumption in Section 6 by predicting passenger itineraries from historical data. Due to data unavailability, the model does not capture crew connections but these could be easily added (like aircraft but unlike passengers, crews induce a one-to-one mapping between flights).

The model determines each flight's departure and arrival times (direct variables), and replicates network-wide passenger flows (indirect variables). We assume that each connecting passenger travels on his/her planned first-leg flight, and gets re-accommodated on a different second-leg flight *if and when* the connection becomes infeasible. We say that a passenger is "accommodated" on a flight if the flight is used as a second leg. Passenger delays include the arrival delays of non-stop passengers, as well as the added travel time resulting from the re-accommodation of passengers traveling on disrupted itineraries. Ultimately, the formulation captures the overall impact of itinerary disruptions across passenger networks, given flight frequencies, aircraft capacities and load factors.

3.2.   (GDP-PAX) Formulation

Inputs

$\mathcal{K}$ = set of airports subject to GDP initiatives, $\{1, \cdots, K\}$

$$\mathcal{T} = \text{set of 15-minute time periods, } \{1, \cdots, T\}$$
$$\mathcal{F} = \text{set of flights, } \{1, \cdots, F\}$$
$$\mathcal{F}_k^a / \mathcal{F}_k^d = \text{subset of flights } i \in \mathcal{F} \text{ scheduled to land / take off at airport } k \in \mathcal{K}$$
$$\underline{T}_i^d / \underline{T}_i^a = \text{earliest time period in } \mathcal{T} \text{ when flight } i \in \mathcal{F} \text{ can depart / arrive}$$
$$\overline{T}_i^d / \overline{T}_i^a = \text{latest time period in } \mathcal{T} \text{ when flight } i \in \mathcal{F} \text{ can depart / arrive}$$
$$\mathcal{C} = \text{set of flight pairs } (i, j) \in \mathcal{F} \times \mathcal{F} \text{ such that an aircraft connects between } i \text{ and } j$$
$$\mathcal{P} = \text{set of passenger itineraries, } \{1, \cdots, P\}$$
$$\mathcal{P}_N / \mathcal{P}_C = \text{subset of non-stop / one-stop passenger itineraries in } \mathcal{P}$$
$$f_1(p) / f_2(p) = \text{first / second flight leg of one-stop passenger itinerary } p \in \mathcal{P}_C$$
$$f_0(p) = \text{flight associated with non-stop passenger itinerary } p \in \mathcal{P}_N$$
$$\mathcal{R}_p = \text{subset of flights } i \in \mathcal{F} \text{ that can accommodate passengers on itinerary } p \in \mathcal{P}$$
$$\mathcal{S}_{kt} = \text{set of segments of the capacity envelope at airport } k \in \mathcal{K} \text{ in period } t \in \mathcal{T}$$
$$C_{it}^d / C_{it}^a = \text{delay cost of flight } i \in \mathcal{F} \text{ if it departs / arrives in period } t \in \mathcal{T}$$
$$\Delta_i^{\min} / \Delta_i^{\max} = \text{minimum / maximum block times of flight } i \in \mathcal{F}$$
$$\tau_{ij} = \text{minimum aircraft turnaround time between flights } i \text{ and } j, \text{ for all } i, j \in \mathcal{C}$$
$$\sigma_{pi} = \text{minimum connecting time between } f_1(p) \text{ and } i \in \mathcal{R}_p \text{ for passengers on itinerary } p \in \mathcal{P}_C$$
$$n_p = \text{number of passengers traveling on itinerary } p \in \mathcal{P}$$
$$\Omega_i = \text{number of passengers that can be accommodated on flight } i \in \mathcal{F}$$
$$\delta_{it}^0 = \text{arrival delay of flight } i \in \mathcal{F} \text{ if it arrives in period } t \in \mathcal{T}$$
$$\delta_{pi}^C = \text{connection delay of passengers on itinerary } p \in \mathcal{P} \text{ if accommodated on flight } i \in \mathcal{R}_i$$
$$\delta_p^S = \text{connection delay of passengers on itinerary } p \in \mathcal{P} \text{ if accommodated on a "sink" option}$$
$$\alpha_{kst}, \beta_{kst}, \gamma_{kst} = \text{parameters defining the capacity envelope at airport } k \in \mathcal{K} \text{ in period } t \in \mathcal{T}$$

The set $\mathcal{K}$ comprises the airports subject to GDP initiatives—all other airports are assumed to be uncapacitated. The set $\mathcal{F}$ includes all flights that are flown by an aircraft or a passenger that visits one of the airports in $\mathcal{K}$. For instance, if an aircraft flies from JFK to MSP to LAX, and a passenger flies from JFK to MSP to SEA, the three flights JFK→MSP, MSP→LAX and MSP→SEA need to be included (even if, for instance, only JFK and LAX are included in $\mathcal{K}$).

Regarding flight operations, we allow for variations in en-route times between $\Delta_i^{\min}$ and $\Delta_i^{\max}$ for each flight i. The gap between $\Delta_i^{\min}$ and $\Delta_i^{\max}$ captures flexibility in flying speeds, aircraft routing, and flying altitude. Next, $\underline{T}_i^d$ is set to flight i's scheduled departure time (flights are not allowed to depart early), but $\underline{T}_i^a$ may not coincide with its scheduled arrival time (due to en-route flexibility). The inputs $\mathcal{C}$ and τ_{ij} track aircraft rotations to maintain aircraft connectivity.

We use the expressions from Bertsimas, Lulli, and Odoni (2011) to compute flight delay costs. Let d_i and a_i denote the periods when flight i is scheduled to depart and arrive, respectively. For each $i \in \mathcal{F}$ and $t \in \mathcal{T}$, we set: $C_{it}^d = (\max(t - d_i, 0))^{1+\varepsilon_1}$ and $C_{it}^a = (\max(t - a_i, 0))^{1+\varepsilon_2}$. We set $\varepsilon_2 > \varepsilon_1$, to reflect that airborne delays are more costly than ground delays (so the model will absorb delays on the ground rather than in the air, whenever possible). The non-linearities ensure inter-flight equity by preventing that disproportionate delays will be assigned to any flight.

We represent airport capacities with convex and piecewise-linear *capacity envelopes*, which capture the non-increasing relationships between arrival and departure throughput (Gilbo 1993). The envelope at airport $k \in \mathcal{K}$ in time period $t \in \mathcal{T}$ is characterized by a set $\mathcal{S}_{kt}$ of linear segments of the form $\alpha_{kts} X_{kt} + \beta_{kts} Y_{kts} \leq \gamma_{kts}$, where X_{kt} and Y_{kt} denote the number of arrivals and departures that can be operated at airport k in period t. The time-dependency reflects variations in weather conditions, runway configurations, and other operating conditions.

The set of passenger itineraries is partitioned between non-stop and one-stop itineraries which, collectively, account for 97.5% of itineraries in the United States (Barnhart, Fearing, and Vaze 2014). Each non-stop itinerary $p \in \mathcal{P}_N$ is booked on exactly one flight $f_0(p)$. We measure the delay of non-stop passengers as their arrival delay, i.e., $\delta_{it}^0 = \max(t - a_i, 0)$ for all $i \in \mathcal{F}$ and $t \in \mathcal{T}$. For each one-stop itinerary $p \in \mathcal{P}_C$, $f_1(p)$ and $f_2(p)$ denote its first-leg and second-leg flights. The set of accommodation options $\mathcal{R}_p$ is defined as the set of flights that can accommodate passengers on itinerary $p \in \mathcal{P}_C$, i.e., flights on the same origin-destination pair as $f_2(p)$, departing after the scheduled arrival time of $f_1(p)$ plus a minimum connecting time, and operated by the same airline or a partner airline. The connection delay δ_{pi}^C is defined as the difference between the second-leg flight's scheduled arrival time and the itinerary's scheduled arrival time, if positive, i.e., $\delta_{pi}^C = \max\left(a_i - a_{f_2(p)}, 0\right)$ for all $p \in \mathcal{P}$ and $i \in \mathcal{R}_p$. Passenger accommodations are constrained by "effective" aircraft capacities Ω_i, equal to the number of aircraft seats minus the number of non-stop passengers on flight i and the number of one-stop passengers with flight i as their first leg (who will always travel on flight i). Finally, we consider a "sink" accommodation to ensure feasibility. It corresponds to an equivalent flight the following day, and it is associated with a high cost δ_p^S.

Decision Variables

$$w_{it}^a / w_{it}^d = \begin{cases} 1 & \text{if flight } i \in \mathcal{F} \text{ is rescheduled to land / take off } by \text{ period } t \in \mathcal{T} \\ 0 & \text{otherwise} \end{cases}$$

$$\lambda_{pi} = \begin{cases} 1 & \text{if the accommodation of passenger itinerary } p \in \mathcal{P}_C \text{ on flight } i \in \mathcal{R}_p \text{ is infeasible} \\ 0 & \text{otherwise} \end{cases}$$

$$z_{pi} = \text{number of passengers on itinerary } p \in \mathcal{P} \text{ accommodated on flight } i \in \mathcal{R}_p$$

$$z_p^S = \text{number of passengers on itinerary } p \in \mathcal{P} \text{ accommodated on the "sink" itinerary}$$

Using notations from Bertsimas and Stock Patterson (1998), each row $w_{i,\cdot}^d$ and $w_{i,\cdot}^a$ takes the form $(0, 0, 0, 1, 1, 1, 1)$. By convention, we assume that $w_{i0}^a = w_{i0}^d = 0$ for each flight i. The variables λ_{pi}, z_{pi} and z_p^S track passenger accommodations—they do not capture actual decisions made by air traffic managers, but quantify the impact of GDP decisions on passenger flows in flight networks.

Mathematical Formulation

$$\min \quad \sum_{i \in \mathcal{F}} \left(\sum_{t \in \mathcal{T}} C_{it}^a \left(w_{it}^a - w_{i,t-1}^a \right) + \sum_{t \in \mathcal{T}} C_{it}^d \left(w_{it}^d - w_{i,t-1}^d \right) \right)$$

$$+ \rho \left[\sum_{p \in \mathcal{P}_N} \sum_{t \in \mathcal{T}} n_p \delta^0_{f_0(p),t} \left(w^a_{f_0(p)t} - w^a_{f_0(p),t-1} \right) + \sum_{p \in \mathcal{P}_C} \left(\sum_{i \in \mathcal{R}_p} \delta^C_{pi} z_{pi} + \delta^S_p z^S_p \right) \right] \tag{3}$$

$$\text{s.t.} \quad w^a_{it} \geq w^a_{i,t-1} \text{ and } w^d_{it} \geq w^d_{i,t-1} \qquad\qquad \forall i \in \mathcal{F}, \forall t \in \mathcal{T} \tag{4}$$

$$w^d_{i,\underline{T}^d_i - 1} = 0, w^a_{i,\underline{T}^a_i - 1} = 0 \qquad\qquad \forall i \in \mathcal{F} \tag{5}$$

$$w^d_{i,\overline{T}^d_i} = 1, w^a_{i,\overline{T}^a_i} = 1 \qquad\qquad \forall i \in \mathcal{F} \tag{6}$$

$$\sum_{t \in \mathcal{T}} w^d_{it} - \sum_{t \in \mathcal{T}} w^a_{it} \geq \Delta^{\min}_i \qquad\qquad \forall i \in \mathcal{F} \tag{7}$$

$$\sum_{t \in \mathcal{T}} w^d_{it} - \sum_{t \in \mathcal{T}} w^a_{it} \leq \Delta^{\max}_i \qquad\qquad \forall i \in \mathcal{F} \tag{8}$$

$$\sum_{t \in \mathcal{T}} w^a_{it} - \sum_{t \in \mathcal{T}} w^d_{jt} \geq \tau_{ij} \qquad\qquad \forall (i,j) \in \mathcal{C} \tag{9}$$

$$\alpha_{kts} \sum_{i \in \mathcal{F}^a_k} \left(w^a_{it} - w^a_{i,t-1} \right) + \beta_{kts} \sum_{i \in \mathcal{F}^d_k} \left(w^d_{it} - w^d_{i,t-1} \right) \leq \gamma_{kts} \qquad \forall k \in \mathcal{K}, \forall t \in \mathcal{T}, \forall s \in \mathcal{S}_{kt} \tag{10}$$

$$\sum_{t \in \mathcal{T}} w^a_{f_1(p),t} - \sum_{t \in \mathcal{T}} w^d_{it} - \sigma_{pi} + M^{(1)}_{pi} \lambda_{pi} \geq 0 \qquad\qquad \forall p \in \mathcal{P}_C, \forall i \in \mathcal{R}_p \tag{11}$$

$$\sum_{t \in \mathcal{T}} w^a_{f_1(p),t} - \sum_{t \in \mathcal{T}} w^d_{f_2(p)t} - \sigma_{p,f_2(p)} - M^{(2)}_p (1 - \lambda_{p,f_2(p)}) + 1 \leq 0 \qquad\qquad \forall p \in \mathcal{P}_C \tag{12}$$

$$z_{pi} \leq n_p(1 - \lambda_{pi}) \qquad\qquad \forall p \in \mathcal{P}_C, \forall i \in \mathcal{R}_p \tag{13}$$

$$n_p(1 - \lambda_{p,f_2(p)}) = z_{p,f_2(p)} \qquad\qquad \forall p \in \mathcal{P}_C \tag{14}$$

$$\sum_{i \in \mathcal{R}_p} z_{pi} + z^S_p = n_p \qquad\qquad \forall p \in \mathcal{P}_C \tag{15}$$

$$\sum_{p \in \mathcal{P}} z_{pi} \leq \Omega_i \qquad\qquad \forall i \in \mathcal{F} \tag{16}$$

$$w^a, w^d \text{ binary} \tag{17}$$

$$\lambda \text{ binary}, z, z^S \text{ integer} \tag{18}$$

Equation (3) optimizes the trade-offs between flight delay costs and passenger delays, weighted by the parameter ρ (when $\rho = 0$, (GDP-PAX) reduces to baseline flight-centric GDP models). Flight delay costs capture arrival and departure delays, weighted by the parameters C^a and C^d. Passenger delays include the arrival delay of non-stop passengers $\sum_{p \in \mathcal{P}_N} \sum_{t \in \mathcal{T}^a_i} n_p \delta^0_{f_0(p),t} \left(w^a_{f_0(p)t} - w^a_{f_0(p),t-1} \right)$, and the added trip times of one-stop passengers $\sum_{p \in \mathcal{P}_C} \left(\sum_{i \in \mathcal{R}_p} \delta^C_{pi} z_{pi} + \delta^S_p z^S_p \right)$. Recall that the connection delay measures the difference between the *scheduled* arrival times between planned vs. realized itineraries. Ideally, we would consider the arrival delay of re-accommodated passengers, by replacing the term δ^C_{pi} by $\delta^C_{pi} + \sum_{t \in \mathcal{T}^a_i} \left(w^a_{it} - w^a_{i,t-1} \right)$. But this would lead to a non-linear objective, thus significantly increasing the model's computational complexity. Since each flight's delay is typically much smaller than the added travel times due to misconnections, this simplification does not alter the structure of the optimal solution—in fact, our results will show that the insights are consistent when considering the added trip times or the arrival delay of one-stop passengers.

Equations (4) to (10) are flight operating constraints. Constraint (4) ensures that w^a and w^d are non-decreasing in t, consistently with their definition. Constraints (5) and (6) state that each flight departs and arrives within the specified time windows. Constraints (7) and (8) impose the minimal and maximal

block-times. Constraint (9) maintains sufficient turnaround times for all aircraft connections. Constraint (10) applies each airport's capacity by restricting the number of arrivals and departures that can be operated in each time period. All these constraints are consistent with the ATFM literature.

Turning to passenger accommodations, Constraint (11) defines the feasibility of connections between first-leg and second-leg flights and Constraint (12) ensures that $\lambda_{p,f_2(p)} = 0$ if the connection between $f_1(p)$ and $f_2(p)$ is feasible. Indeed, $\sum_{t \in \mathcal{T}} w^a_{f_1(p),t} - \sum_{t \in \mathcal{T}} w^d_{it}$ characterizes the connecting time between the first-leg flight $f_1(p)$ and flight i. To ensure that Constraint (11) is satisfied whenever $\lambda_{pi} = 1$ and that Constraint (12) is satisfied whenever $\lambda_{p,f_2(p)} = 0$, we define the scalars $M^{(1)}_{pi} = \sigma_{pi} + \overline{T}^a_{f_1(p)} - \underline{T}^d_i$ for all $p \in \mathcal{P}_C$ and all $i \in \mathcal{F}$, and $M^{(2)}_p = \sigma_{p,f_2(p)} + \underline{T}^a_{f_1(p)} - \overline{T}^d_{f_2(p)}$ for all $p \in \mathcal{P}_C$. Proposition 4 (proved in Appendix B) shows that Constraints (11) and (12) together ensure that $\lambda_{pi} = 1$ if the connecting time is lower than the minimum connecting time σ_{pi}, and 0 otherwise. It also shows that Constraint (12) is critical to obtain this result.

PROPOSITION 4. *There exists an optimal solution to (GDP-PAX) such that $\lambda_{pi} = 0$ for each $p \in \mathcal{P}_C, i \in \mathcal{R}_p$ that satisfies $\sum_{t \in \mathcal{T}} w^a_{f_1(p),t} - \sum_{t \in \mathcal{T}} w^d_{it} \geq \sigma_{pi}$. This property is not satisfied if Constraint (12) is omitted.*

Next, Constraint (13) ensures that no passenger gets assigned to a second-leg flight if the connection is not feasible. In airline disruption management, Bratu and Barnhart (2006) and (Marla, Vaaben, and Barnhart 2017) use variables $\tilde{\lambda}_p = 1$ if itinerary $p \in \mathcal{P}_C$ is disrupted, and 0 otherwise, and $\tilde{z}_{pr}$ equal to the number of passengers re-accommodated from itinerary $p \in \mathcal{P}_C$ to itinerary $r \in \mathcal{P}_C$. They write constraints of the form $\sum_{p \in \mathcal{P}_C} \tilde{z}_{pr} \leq \left(\sum_{p \in \mathcal{P}_C} n_p \right) (1 - \tilde{\lambda}_r), \forall r \in \mathcal{P}_C$. Constraint (13) improves them in two ways. First, Proposition 5 (proved in Appendix B) shows that the constraints $\tilde{z}_{pr} \leq n_p(1 - \tilde{\lambda}_r), \forall p, r \in \mathcal{P}_C$ are valid inequalities with a tighter linear programming relaxation. Second, we use our assumption that all passengers travel on their original first-leg flight to re-formulate these constraints using matrices of size $|\mathcal{P}| \times |\mathcal{F}|$ rather than $|\mathcal{P}|^2$.

PROPOSITION 5. *Let us consider the variables $\tilde{\lambda}_p = 1$ if itinerary $p \in \mathcal{P}_C$ is disrupted, and 0 otherwise, and $\tilde{z}_{pr}$ equal to the number of passengers originally planned on itinerary $p \in \mathcal{P}_C$ and re-accommodated on itinerary $r \in \mathcal{P}_C$. Then, if $\tilde{z}_{pr} \leq n_p$ for all $p \in \mathcal{P}_C$, the constraint $\tilde{z}_{pr} \leq n_p(1 - \tilde{\lambda}_r), \forall p, r \in \mathcal{P}_C$ is a valid inequality that yields a tighter linear programming relaxation than the constraint $\sum_{p \in \mathcal{P}_C} \tilde{z}_{pr} \leq \left(\sum_{p \in \mathcal{P}_C} n_p \right) (1 - \tilde{\lambda}_r), \forall r \in \mathcal{P}_C$.*

The remaining constraints ensure that passengers get accommodated on their planned itineraries if they are not disrupted (Equation (14)), that all passengers get accommodated on one flight or the "sink" option (Equation (15)), and that passenger accommodations meet effective aircraft capacities (Equation (16)). Constraints (17) and (18) define the domain of definition of the variables.

3.3. Size of the Formulation

We compare our (GDP-PAX) model to a baseline flight-centric GDP model, referred to as (GDP), obtained by considering only the decision variables w^a and w^d and the flight operating constraints (Equations (4) to (10)). Table 1 reports the number of variables and constraints of both models, where $|\mathcal{P}_C|$ denotes the number of one-stop passenger itineraries and $\bar{S}$ denotes an upper bound on the number of segments in each airport's capacity envelope. The size of (GDP-PAX) scales up with the number of itinerary-flight pairs, resulting in $\mathcal{O}(|\mathcal{P}_C| F)$ variables and constraints. In theory, $|\mathcal{P}_C|$ scales up quadratically with the number of

flights F. However, the set of passenger itineraries is sparse because feasible itineraries can only be created by a handful of flight pairs in $\mathcal{F} \times \mathcal{F}$ (i.e., (i,j) pairs such that the departure airport j is the same as the arrival airport of i, with sufficient connecting time, and flown by the same airline or partner airlines). Similarly, for each one-stop itinerary $p \in \mathcal{P}_C$, the only relevant variables are related to the flights in the subset $\mathcal{R}_p$ rather than in the full set $\mathcal{F}$. This set is also sparse, for the same reasons. The number of non-zero variables and constraints can thus be ultimately significantly lower than F^3, or even than $|\mathcal{P}_C| F$.

Table 1   Size of the (GDP) and (GDP-PAX) models.

Metric	(GDP)	(GDP-PAX)
# binary variables	$2FT$	$2FT + \mathcal{O}\left(|\mathcal{P}_C| F\right)$
# integer variables	–	$\mathcal{O}\left(|\mathcal{P}_C| F\right) + |\mathcal{P}_C|$
# constraints (upper bound)	$2FT + 4F + F^2 + \bar{S}KT$	$2FT + 5F + F^2 + \bar{S}KT + 2|\mathcal{P}_C| F + 2|\mathcal{P}_C|$

4.   Solution Algorithm

Any realistic GDP model results in large-scale optimization problems, with huge memory and solution time requirements. We adopt a rolling approach to solve the model iteratively for a limited look-ahead window. This approach is consistent with practice and the recent literature (e.g. Bertsimas, Lulli, and Odoni 2011). Our passenger-centric approach complicates this rolling procedure, however, because we need to link flight operations and passenger itineraries across the full day to maintain feasibility in passenger flows and accurately estimate passenger delays. We describe next our algorithmic procedure and the adjustments to the (GDP-PAX) formulation that make it amenable to implementation in a rolling horizon.

We apply (GDP-PAX) in successive decision epochs, indexed by $u = 1, \cdots, U$ for a look-ahead window of length w (set to 4, 5 or 6 hours). At each epoch u, we denote by $\mathcal{T}_u$ the set of time periods corresponding to the interval $[u, u+w]$, plus additional "buffer" periods to maintain feasibility in case some flights need to be delayed beyond the look-ahead window. We denote by $\widetilde{d}_i$ and $\widetilde{a}_i$ the planned departure and arrival times of flight $i \in \mathcal{F}$ at time u (obtained from the schedule of flights or the model's output in the preceding decision epochs). Let $\widetilde{\mathcal{F}}^d = \left\{ i \in \mathcal{F} \mid u \leq \widetilde{d}_i \leq u + w \right\}$ and $\widetilde{\mathcal{F}}^a = \left\{ i \in \mathcal{F} \mid u \leq \widetilde{a}_i \leq u + w \right\}$ track the flights planned to depart and arrive, respectively, during the look-ahead window. We include in the model the set $\mathcal{F}$ of all flights yet to arrive across the full day, in order to track passenger flows beyond the look-ahead window $[u, u+w]$. However, the flight decision variables w^d and w^a are restricted to the sets $\widetilde{\mathcal{F}}^d$ and $\widetilde{\mathcal{F}}^a$, respectively.

At each decision epoch, we include all one-stop itineraries subject to the flight operating decisions within the look-ahead window $[u, u+w]$, and whose passengers have not all been provided final accommodations. We define *final accommodations* through parameters ζ_{pi} (resp. ζ_p^S) that denote the number of passengers booked on itinerary $p \in \mathcal{P}_C$, accommodated on flight $i \in \mathcal{R}_p$ (resp. on the sink option) and who can no longer be provided any other accommodation. Specifically, ζ_{pi} takes the value z_{pi} (obtained from the model at the preceding epoch) if flight i departed before time u (in which case the passenger is already en-route to the destination), or if flight i corresponds to the second-leg flight $f_2(p)$ and passengers will make their connection with certainty. The latter condition occurs if the first-leg flight $f_1(p)$ already arrived and sufficient connecting

time is available (even if $f_2(p)$ is not delayed), or if $f_1(p)$ already departed and sufficient connecting time will be available (even if the en-route time of $f_1(p)$ takes its maximal value and $f_2(p)$ is not delayed). Similarly, ζ_p^S takes the value z_p^S if the accommodation of all passengers on itinerary p is final, or if all the candidate second-leg flights have departed already. The logic is summarized in Equations (19) and (20).

$$\zeta_{pi} = \begin{cases} z_{pi} & \text{if } \widetilde{d}_i < u \\ z_{pi} & \text{if } i = f_2(p), \ \widetilde{a}_{f_1(p)} < u \text{ and } \widetilde{a}_{f_1(p)} + \sigma_{pi} \leq \widetilde{d}_i \\ z_{pi} & \text{if } i = f_2(p), \ \widetilde{d}_{f_1(p)} < u \text{ and } \widetilde{d}_{f_1(p)} + \sigma_{pi} + \Delta_i^{\max} \leq \widetilde{d}_i \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

$$\zeta_p^S = \begin{cases} z_p^S & \text{if } \sum_{i \in \mathcal{F}} \zeta_{pi} = n_p, \text{ or if } \widetilde{d}_i < u \text{ for all } i \in \mathcal{R}_p \\ 0 & \text{otherwise} \end{cases} \tag{20}$$

The subset of one-stop passenger itineraries included in the model, denoted by $\widetilde{\mathcal{P}}_C$, is then given by:

$$\widetilde{\mathcal{P}}_C = \left\{ p \in \mathcal{P}_C \mid \widetilde{d}_{f_1(p)} \leq u + w \text{ and } \sum_{i \in \mathcal{F}} \zeta_{pi} + \zeta_p^S < n_p \right\} \tag{21}$$

We then update the parameters n_p and Ω_i by the number of passengers booked on itinerary $p \in \widetilde{\mathcal{P}}_C$ that remain to be accommodated and the remaining seating capacity on the aircraft used to fly flight $i \in \mathcal{F}$:

$$n_p \leftarrow n_p - \sum_{i \in \mathcal{F}} \zeta_{pi} - \zeta_p^S \qquad \forall p \in \widetilde{\mathcal{P}}_C \tag{22}$$

$$\Omega_i \leftarrow \Omega_i - \sum_{p \in \mathcal{P}_C} \zeta_{pi} \qquad \forall i \in \mathcal{F} \tag{23}$$

We can implement the flight operating constraints (Equations (4) to (10)), restricted to time periods in $\mathcal{T}_u$ and flights in $\widetilde{\mathcal{F}}^a$ and $\widetilde{\mathcal{F}}^d$. We can also implement the passenger flow constraints (Equations (13) to (16)), using one-stop itineraries in $\widetilde{\mathcal{P}}_C$ and all flights in $\mathcal{F}$. But we need to redefine the passenger accommodation constraints (Equations (11) and (12)) because the first-leg and second-leg flights of each itinerary in $\widetilde{\mathcal{P}}_C$ are not necessarily included in $\widetilde{\mathcal{F}}^a$ and $\widetilde{\mathcal{F}}^d$. To this end, we partition the set $\widetilde{\mathcal{P}}_C$ into three subsets, denoted by $\widetilde{\mathcal{P}}_C^A$, $\widetilde{\mathcal{P}}_C^D$, and $\widetilde{\mathcal{P}}_C^0$ (Equations (24) to (26)), and formulate the related constraints in Equations (27) to (33).

- Subset $\widetilde{\mathcal{P}}_C^A$ includes itineraries whose first-leg flights are scheduled to arrive within the look-ahead window. For each flight $i \in \mathcal{R}_p \cap \mathcal{F}_t^d$ (scheduled to depart during the look-ahead window), Equation (27) is identical to Equation (11). Similarly, if $f_2(p) \in \mathcal{F}_t^d$, Equation (28) is identical to Equation (12). Then, for each flight $i \in \mathcal{R}_p \setminus \mathcal{F}_t^d$, Equation (29) states that passengers on itinerary p can be accommodated on flight i if and only if $f_1(p)$ arrives by $\widetilde{d}_i - \sigma_{pi}$ (the latest arrival time to enable passenger connections).

- Subset $\widetilde{\mathcal{P}}_C^D$ includes itineraries whose first-leg flights are scheduled to depart within the look-ahead window, but to arrive after $u + w$. For each flight $i \in \mathcal{R}_p \cap \mathcal{F}_t^d$, Equations (30) and (31) are analogous to Equations (11) and (12), except that the arrival time of flight $f_1(p)$ is not controlled. We thus state that accommodations on flight i are infeasible if the connecting time is insufficient even if the en-route time of $f_1(p)$ takes its minimal value $\Delta_{f_1(p)}^{\min}$. Then, for each flight $i \in \mathcal{R}_p \setminus \mathcal{F}_t^d$, Equation (32) states that passengers on itinerary p can be accommodated on flight i if and only if $f_1(p)$ departs by $\widetilde{d}_i - \sigma_{pi} - \Delta_{f_1(p)}^{\min}$ (the latest departure time to enable passenger connections if the en-route time takes its minimal value).

- Subset $\widetilde{\mathcal{P}}_C^0$ includes itineraries whose first-leg flights are not controlled by the model at epoch u, either because $f_1(p)$ leaves before u and arrives after $u+w$, or because $f_1(p)$ already arrived by time u. For each flight $i \in \mathcal{R}_p \cap \mathcal{F}_t^d$, Equation (33) states that passengers can be accommodated on flight i if and only if flight i departs no earlier than $\widetilde{a}_{f_1(p)} + \sigma_{p,f_2(p)}$ (the earliest departure time to enable passenger connections). For each flight $i \in \mathcal{R}_p \setminus \mathcal{F}_t^d$, we enforce the constraints by fixing the variables λ_{pi}.

Note that Equations (29), (32) and (33) capture that $\lambda_{pf_2(p)} = 1$ *if and only if* the connecting time between $f_1(p)$ and $f_2(p)$ is insufficient, so there is no need for a constraint analogous to Equation (12) in these cases.

$$\widetilde{\mathcal{P}}_C^A = \left\{ p \in \widetilde{\mathcal{P}}_C \mid f_1(p) \in \widetilde{\mathcal{F}}^a \right\} \tag{24}$$

$$\widetilde{\mathcal{P}}_C^D = \left\{ p \in \widetilde{\mathcal{P}}_C \mid f_1(p) \in \widetilde{\mathcal{F}}^d \setminus \widetilde{\mathcal{F}}^a \right\} \tag{25}$$

$$\widetilde{\mathcal{P}}_C^0 = \left\{ p \in \widetilde{\mathcal{P}}_C \mid f_1(p) \notin \left(\widetilde{\mathcal{F}}^a \cup \widetilde{\mathcal{F}}^d \right) \right\} \tag{26}$$

$$\sum_{t \in \mathcal{T}} w_{f_1(p),t}^a - \sum_{t \in \mathcal{T}} w_{it}^d - \sigma_{pi} + M_{pi}^{(1)} \lambda_{pi} \geq 0 \qquad \forall p \in \widetilde{\mathcal{P}}_C^A, \forall i \in \mathcal{R}_p \cap \widetilde{\mathcal{F}}^d \tag{27}$$

$$\sum_{t \in \mathcal{T}} w_{f_1(p),t}^a - \sum_{t \in \mathcal{T}} w_{f_2(p)t}^d - \sigma_{p,f_2(p)} - M_p^{(2)}(1 - \lambda_{p,f_2(p)}) + 1 \leq 0 \qquad \forall p \in \widetilde{\mathcal{P}}_C^A \mid f_2(p) \in \widetilde{\mathcal{F}}^d \tag{28}$$

$$\lambda_{pi} = 1 - w_{f_1(p), \tilde{d}_i - \sigma_{pi}}^a \qquad \forall p \in \widetilde{\mathcal{P}}_C^A, \forall i \in \mathcal{R}_p \setminus \widetilde{\mathcal{F}}^d \tag{29}$$

$$\sum_{t \in \mathcal{T}} w_{f_1(p),t}^d - \sum_{t \in \mathcal{T}} w_{it}^d - \sigma_{pi} - \Delta_{f_1(p)}^{\min} + M_{pi}^{(1)} \lambda_{pi} \geq 0 \qquad \forall p \in \widetilde{\mathcal{P}}_C^D, \forall i \in \mathcal{R}_p \cap \widetilde{\mathcal{F}}^d \tag{30}$$

$$\sum_{t \in \mathcal{T}} w_{f_1(p),t}^d - \sum_{t \in \mathcal{T}} w_{f_2(p)t}^d - \sigma_{p,f_2(p)} - \Delta_{f_1(p)}^{\min} - M_p^{(2)}(1 - \lambda_{p,f_2(p)}) + 1 \leq 0 \qquad \forall p \in \widetilde{\mathcal{P}}_C^D \mid f_2(p) \in \widetilde{\mathcal{F}}^d \tag{31}$$

$$\lambda_{pi} = 1 - w_{f_1(p), \tilde{d}_i - \sigma_{pi} - \Delta_{f_1(p)}^{\min}}^d \qquad \forall p \in \widetilde{\mathcal{P}}_C^D, \forall i \in \mathcal{R}_p \setminus \widetilde{\mathcal{F}}^d \tag{32}$$

$$\lambda_{pi} = w_{i, \tilde{a}_{f_1(p)} + \sigma_{p,f_2(p)} - 1}^d \qquad \forall p \in \widetilde{\mathcal{P}}_C^0, \forall i \in \mathcal{R}_p \cap \widetilde{\mathcal{F}}^d \tag{33}$$

Algorithm 1 summarizes our full rolling procedure to solve (GDP-PAX).

Algorithm 1 Rolling horizon algorithm to solve (GDP-PAX).

– Inputs: flight schedule, passenger itineraries, aircraft capacities airport capacities

– Initialize final itineraries $\zeta_{pi} = 0, \forall p \in \mathcal{P}_C, i \in \mathcal{R}_p$, and $\zeta_p^S = 0, \forall p \in \mathcal{P}_C$

– Define rolling periods $u = 1, \cdots, U$ and duration of each look-ahead window w

    for $u = 1, \cdots, U$ **do**

      – Define flight sets: $\widetilde{\mathcal{F}}^a = \{ i \in \mathcal{F} \mid u \leq \widetilde{a}_i \leq u + w \}$ and $\widetilde{\mathcal{F}}^d = \left\{ i \in \mathcal{F} \mid u \leq \tilde{d}_i \leq u + w \right\}$

      – Define passenger sets: $\widetilde{\mathcal{P}}_C = \left\{ p \in \mathcal{P}_C \mid \tilde{d}_{f_1(p)} \leq u + w \text{ and } \sum_{i \in \mathcal{F}} \zeta_{pi} + \zeta_p^S < n_p \right\}$, and subsets $\widetilde{\mathcal{P}}_C^A, \widetilde{\mathcal{P}}_C^D$, and $\widetilde{\mathcal{P}}_C^0$ (Equations (24) to (26))

      – Solve the (GDP-PAX) model: Equations (3) to (10), Equations (27) to (31), and Equations (13) to (18)

      – Update the arrival schedule $\widetilde{a}_i$ for $i \in \widetilde{\mathcal{F}}^a$, and the departure schedule $\tilde{d}_i$ for $i \in \widetilde{\mathcal{F}}^d$

      – Fix final itineraries; update parameters ζ_{pi} for $p \in \mathcal{P}_C, i \in \mathcal{R}_p$, and ζ_p^S for $p \in \mathcal{P}_C$ (Equations (19) and (20))

      – Update the number of remaining passengers and effective aircraft capacities (Equations (22) and (23))

    end for

We could solve the baseline (GDP) model with Algorithm 1, by setting $\rho = 0$. But this approach would artificially increase the computational complexity of (GDP) by computing passenger flows—hence, underestimate the computational costs of (GDP-PAX) in comparison. Moreover, it would not necessarily provide optimal passenger flows, given the flight-level decisions w^a and w^d—hence, overestimate the benefits of (GDP-PAX) in comparison. To ensure a fair assessment of (GDP-PAX), we implement Algorithm 2, which first solves (GDP) without passenger-centric constraints and then optimizes passenger accommodations (by solving a passenger-flow model equivalent to (GDP-PAX) where the variables w^a and w^d are fixed).

Algorithm 2 Rolling horizon algorithm to solve (GDP) with separate passenger flow model.

– Inputs: flight schedule, passenger itineraries, aircraft capacities and airport capacities

– Define rolling periods $u = 1, \cdots, U$ and duration of each look-ahead window w

   for $u = 1, \cdots, U$ **do**

      – Define flight sets: $\widetilde{\mathcal{F}}^a = \left\{ i \in \mathcal{F} \mid u \le \widetilde{a}_i \le u + w \right\}$ and $\widetilde{\mathcal{F}}^d = \left\{ i \in \mathcal{F} \mid u \le \widetilde{d}_i \le u + w \right\}$

      – Define passenger sets: $\widetilde{\mathcal{P}}_C = \left\{ p \in \mathcal{P}_C \mid \widetilde{d}_{f_1(p)} \le u + w \text{ and } \sum_{i \in \mathcal{F}} \zeta_{pi} + \zeta_p^S < n_p \right\}$, and subsets $\widetilde{\mathcal{P}}_C^A$, $\widetilde{\mathcal{P}}_C^D$, and $\widetilde{\mathcal{P}}_C^0$ (Equations (24) to (26))

      – Solve the (GDP) model:

$$\min_{w^a, w^d} \sum_{i \in \mathcal{F}} \left(\sum_{t \in \mathcal{T}} C_{it}^a \left(w_{it}^a - w_{i,t-1}^a \right) + \sum_{t \in \mathcal{T}} C_{it}^d \left(w_{it}^d - w_{i,t-1}^d \right) \right) \quad \text{s.t.: Eq. (4)–(10), (17)}$$

      – Update the arrival schedule $\widetilde{a}_i$ for $i \in \widetilde{\mathcal{F}}^a$, and the departure schedule $\widetilde{d}_i$ for $i \in \widetilde{\mathcal{F}}^d$

      – Solve a passenger-flow model to reconstruct passenger itineraries:

$$\min_{\lambda, z, z^S} \sum_{p \in \mathcal{P}_C} \left(\sum_{i \in \mathcal{R}_p} \delta_{pi}^C z_{pi} + \delta_p^S z_p^S \right) \quad \text{s.t.: Eq. (27)–(31), (13)–(16), (18)}$$

      – Fix final itineraries; update parameters ζ_{pi} for $p \in \mathcal{P}_C, i \in \mathcal{R}_p$, and ζ_p^S for $p \in \mathcal{P}_C$ (Equations (19) and (20))

      – Update the number of remaining passengers and effective aircraft capacities (Equations (22) and (23))

   end for

---

5. Experimental Results

We implement (GDP-PAX) using real-world data. Throughout this section, we report results from the full rolling horizon algorithm described in Section 4, thus capturing the overall performance of our model and algorithm over a full day of GDP operations—as opposed to a single iteration at a point in time.

5.1. Experimental Setup

We use data from the US National Aviation System from 2007. This year was chosen because it was one of the busiest years in aviation, and because of the availability of original and reconstructed data (e.g., passenger itineraries, airport capacities). Given the low variability in flight schedules from one day to another, we use scheduling inputs for one representative day, September 18, 2007. The set of airports $\mathcal{K}$ includes up to all the "CORE 30" airports—the 30 most busiest US airports. We first consider instances where the 6 most congested airports (i.e., JFK, EWR, LGA, ATL, ORD and PHL) are subject to GDP initiatives and thus

included in $\mathcal{K}$; we then consider 18-airport GDP instances and instances with all 30 airports. Ultimately, this setting captures GDP instances of the size of the full US and European networks—the largest ones encountered in practice. We vary the look-ahead window w from 4 hours to 6 hours.

Data on flight schedules, aircraft connections and fleet types are obtained from the Aviation System Performance Metrics (ASPM) database, maintained by the Federal Aviation Administration (2013). We set the upper bound of arrival delays to 6 hours. We use flight delay coefficients of $\varepsilon_1 = 1.05$ and $\varepsilon_2 = 1.1$. We assume that each flight has a 10% block-time flexibility (e.g., if a flight is scheduled with a 120-minute block time, its actual block time may vary between 108 and 132 minutes). We use the minimum aircraft turnaround times estimated by Pyrgiotis (2011). We obtain passenger itineraries from Barnhart, Fearing, and Vaze (2014). We consider a minimum passenger connecting time of 30 minutes at each airport, which corresponds to the 5$^{\text{th}}$ percentile of the distribution of planned connection times.

We define airport capacity envelopes in Visual Meteorological Conditions (VMC) and Instrument Meteorological Conditions (IMC)—proxies for "good" and "poor" weather, respectively. At the airports in the New York area (JFK, EWR and LGA), we use the envelopes estimated empirically by Simaiakis (2012) and the pattern of runway configuration utilization from Jacquillat, Odoni, and Webster (2017). At the other 27 airports, we approximate the VMC and IMC capacity envelopes with three segments capturing the arrival, departure and total capacities in 2007 estimated by the Federal Aviation Administration (2004).

We consider two extreme weather scenarios by assuming that weather conditions are either good or bad at all airports in all time periods. Of course, most instances fall between these extremes. But these two scenarios enable to characterize passenger-level GDP under moderate and strong congestion, respectively.

We implement the models using CPLEX 12.7 on an Intel(R) Core(TM) i5 running at 3.1 GHz 16 GB RAM. We look for solutions within an optimality gap of 1% with a one-hour limit.

5.2. Computational Performance

Figure 2 shows the number of flights and passenger itineraries included in the model in each rolling period, in the largest setting (i.e., 30 airports in $\mathcal{K}$ and $w = 6$ hours). Note that (GDP-PAX) involves very large-scale optimization problems, even with the rolling horizon algorithm. At each period, the model optimizes the arrival and/or departure times of up to 14,000 flights and the accommodation of passengers booked on over 60,000 itineraries—resulting in millions of decision variables and constraints.

Table 2 reports the computational performance of (GDP-PAX), with 6, 18 and 30 airports in $\mathcal{K}$ and a look-ahead window $w = 4$, 5 and 6 hours. The "total cost" metric refers to the model's objective (Equation (3)) obtained from the full rolling algorithm. The table also shows the minimum, median and maximum runtimes and the maximal optimality gap across all integer programs solved in the 13–15 iterations of the algorithm.

In all but two instances, (GDP-PAX) derives solutions within the 1% optimality gap. With 6 airports in $\mathcal{K}$, the median runtime is on the order of one minute. Obviously, as the number of airports in $\mathcal{K}$ increases, so do computational times. Nonetheless, median computational times remain reasonable, on the order of 1–3 minutes—enabling the real-time implementation of the model in practice.

Moreover, longer look-ahead results in larger problems and longer computational times. Increasing the look-ahead window from 4 to 6 hours increases median runtimes moderately but increases maximal runtimes
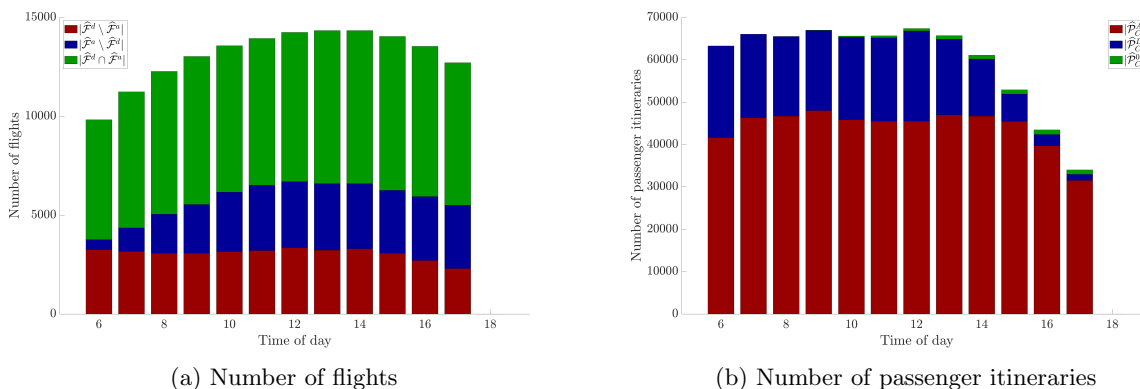
(a) Number of flights

(b) Number of passenger itineraries

Figure 2   Size of the model in each iteration of the rolling horizon algorithm.

more significantly (which can be alleviated by setting a larger optimality tolerance, say 2-3%). At the same time, longer look-ahead windows do not necessarily change the quality of the solution dramatically. In some cases, longer look-ahead even leads to slightly *higher* total costs, due to the approximate rolling horizon algorithm and to variations in solution qualities within the 1% optimality gap. Longer look-ahead is most beneficial for larger values of ρ, i.e., when passenger delays become more prominent—thus underscoring the ripple effects of flight operating decisions on passenger itineraries over the full day of operations.

Most importantly, passenger-centric considerations result in moderate increases in computational times, as compared to the baseline flight-centric model. Despite larger model instances, the runtimes of (GDP-PAX) remain of the same order of magnitude as those of (GDP). For example, with 18 airports, the median runtimes increase from 41-78 seconds to 68-147 seconds. In other words, considering passenger accommodations does not make the GDP problem intractable. Ultimately, this suggests that passenger-centric considerations can be incorporated into existing ATFM decision support systems at limited computational costs.

5.3.   Impact of Passenger-centric GDP Operations

We now evaluate the performance of (GDP-PAX) in terms of solution quality. Table 3 reports flight delay costs and the delays of non-stop and one-stop passengers, for different values of ρ.

The most striking observation is that (GDP-PAX) yields very large reductions in passenger delays at comparatively small increases in flight delay costs, as compared to the (GDP) baseline. As expected, higher values of ρ generally induce lower passenger delays and higher flight delay costs (a few exceptions are due to the rolling approximation and variations of the solutions within the 1% optimality gap). Most critically, the percent-wise reduction in passenger delays far outweighs any increase in flight delay costs across all test instances. In VMC, passenger delays can be reduced by up to 50%–90% by increasing flight delay costs by 3%. In IMC, passenger delay reductions amount to 20% by increasing flight delay costs by 1%, and to 60% by increasing flight delay costs by 3%. For large values of ρ, passenger delays can be cut even more, obviously through higher increases in flight delay costs (by 10% in VMC and 35%–40% in IMC).

Additional insights can be derived from these results. First, non-stop passenger delays can decrease significantly, even when flight delays go up. That is, the choice of *which* flights to delay—by prioritizing flights carrying more passengers—can have significant impacts on non-stop passenger delays. Second, the mechanisms of passenger delay mitigation vary with the extent of congestion. Under moderate congestion (VMC),

Table 2   Computational results.

Weather	w	Metric	(GDP) $\|\mathcal{K}=6\|$	(GDP) $\|\mathcal{K}=18\|$	(GDP) $\|\mathcal{K}=30\|$	(GDP-PAX)—$\rho=0.0001$ $\|\mathcal{K}=6\|$	(GDP-PAX)—$\rho=0.0001$ $\|\mathcal{K}=18\|$	(GDP-PAX)—$\rho=0.0001$ $\|\mathcal{K}=30\|$	(GDP-PAX)—$\rho=0.01$ $\|\mathcal{K}=6\|$	(GDP-PAX)—$\rho=0.01$ $\|\mathcal{K}=18\|$	(GDP-PAX)—$\rho=0.01$ $\|\mathcal{K}=30\|$
VMC	4 hrs.	Total cost	1,836 (base)	2,378 (base)	3,006 (base)	1,877 (base)	2,355 (base)	2,936 (base)	2,241 (base)	2.827 (base)	3,568 (base)
		Min. CPU (s)	13	21	39	14	27	39	15	29	41
		Med. CPU (s)	23	41	53	35	68	82	37	73	83
		Max. CPU (s)	64	105	510	132	227	319	165	287	241
		Max. Gap	<1%	<1%	<1%	<1%	<1%	<1%	<1%	<1%	<1%
VMC	5 hrs.	Total cost	1,900 (+3.5%)	2,300 (-3.3%)	2,915 (-3.0%)	1,874 (-0.1%)	2,383 (+1.2%)	2,928 (-0.3%)	2,105 (-6.0%)	2,678 (-5.3%)	3,446 (-3.4%)
		Min. CPU (s)	16	26	45	20	36	56	36	36	57
		Med. CPU (s)	28	52	71	55	93	119	90	90	116
		Max. CPU (s)	111	516	225	1,439	382	426	3,600	2,232	463
		Max. Gap	<1%	<1%	<1%	<1%	<1%	<1%	1.4%	<1%	<1%
VMC	6 hrs.	Total cost	1,853 (+0.9%)	2,357 (-0.9%)	2,914 (-3.1%)	1,814 (-3.3%)	2,340 (-0.6%)	2,907 (-1.0%)	2,065 (-7.8%)	2,699 (-4.5%)	3,451 (-3.3%)
		Min. CPU (s)	15	33	51	22	51	75	24	56	74
		Med. CPU (s)	32	75	84	57	124	157	57	113	154
		Max. CPU (s)	119	227	376	331	630	554	3,600	630	784
		Max. Gap	<1%	<1%	<1%	<1%	<1%	<1%	1.2%	<1%	<1%
IMC	4 hrs.	Total cost	35,579 (base)	39,973 (base)	42,203 (base)	37,214 (base)	41,631 (base)	43,650 (base)	58,395 (base)	65,029 (base)	69,307 (base)
		Min. CPU (s)	15	30	36	17	34	58	18	32	52
		Med. CPU (s)	21	43	58	37	76	100	41	76	94
		Max. CPU (s)	31	73	107	51	129	184	387	268	393
		Max. Gap	<1%	<1%	<1%	<1%	<1%	<1%	<1%	<1%	<1%
IMC	5 hrs.	Total cost	35,844 (+0.7%)	39,867 (-0.3%)	42,237 (+0.1%)	37,291 (+0.2%)	41,605 (-0.1%)	43,814 (+0.4%)	57,386 (-1.7%)	64,880 (-0.2%)	70,736 (+2.1%)
		Min. CPU (s)	17	34	46	22	45	72	21	47	70
		Med. CPU (s)	29	57	73	51	102	154	53	106	134
		Max. CPU (s)	135	109	198	91	182	330	274	409	547
		Max. Gap	<1%	<1%	<1%	<1%	<1%	<1%	<1%	<1%	<1%
IMC	6 hrs.	Total cost	35,343 (-0.7%)	39,434 (-1.3%)	41,648 (-1.3%)	36,728 (-1.3%)	41,039 (-1.4%)	43,424 (-0.5%)	54,462 (-6.7%)	61,920 (-4.8%)	67,283 (-2.9%)
		Min. CPU (s)	21	42	68	47	62	92	27	67	93
		Med. CPU (s)	33	78	105	92	147	150	74	145	174
		Max. CPU (s)	64	133	246	206	311	687	552	1,075	2,751
		Max. Gap	<1%	<1%	<1%	<1%	<1%	<1%	<1%	<1%	<1%

non-stop passenger delays and one-stop passenger delays both go down significantly. Under strong congestion (IMC), tackling non-stop passenger delays is much harder: there is less flexibility in allocating flight delays to reduce non-stop passenger delays so most benefits stem from avoiding misconnections.

Recall that, to retain linearity, our model estimates the delay of re-accommodated one-stop passengers as the difference between the scheduled time of the new second-leg flight and the original second-leg flight (as opposed to their actual arrival times). To justify this simplification, we report in Table 4 (i) the added scheduled trip times due to misconnections ("misconn"), i.e., $\sum_{p\in\mathcal{P}_C}\sum_{i\in\mathcal{R}_p}\delta_{pi}^C z_{pi}$; (ii) the added arrival delays ("arr. delay"), i.e., $\sum_{p\in\mathcal{P}_C}\sum_{i\in\mathcal{R}_p}\left(\sum_{t\in\mathcal{T}_i^a}\left(w_{it}^a-w_{i,t-1}^a\right)\right)z_{pi}$; and (iii) the costs of re-accommodations to the sink option ("sink"), i.e., $\sum_{p\in\mathcal{P}_C}\delta_p^S z_p^S$. Accordingly, the term in the objective function ("obj.") is

**Table 3 Flight delay costs and passenger delays ($w = 5$ hours).**

Size	Model	VMC				IMC			
		Flight	Non-stop	One-stop	All PAX	Flight	Non-stop	One-stop	All PAX
6 AP	(GDP)	1,900	20,534	225,308	245,842	35,844	874,404	6,250,435	7,124,839
	$\rho = 0.00001$	1,844	13,614	82,804	96,418	36,117	867,260	4,711,420	5,578,680
		(-3.0%)	(-33.7%)	(-63.2%)	(-60.8%)	(+0.8%)	(-0.8%)	(-24.6%)	(-21.7%)
	$\rho = 0.0001$	1,869	12,982	37,674	50,656	37,030	852,925	1,762,994	2,615,919
		(-1.6%)	(-36.8%)	(-83.3%)	(-79.4%)	(+3.3%)	(-2.5%)	(-71.8%)	(-63.3%)
	$\rho = 0.001$	1,902	12,929	11,939	24,868	39,115	783,183	615,535	1,398,718
		(+0.1%)	(-37.0%)	(-94.7%)	(-89.9%)	(+9.1%)	(-10.4%)	(-90.2%)	(-80.4%)
	$\rho = 0.01$	1,994	9,994	1,102	11,096	49,030	641,356	194,262	835,618
		(+5.0%)	(-51.3%)	(-99.5%)	(-95.5%)	(+36.8%)	(-26.7%)	(-96.9%)	(-88.3%)
18 AP	(GDP)	2,300	24,335	238,536	262,871	39,867	982,036	7,412,038	8,394,074
	$\rho = 0.00001$	2,364	19,687	117,646	137,333	40,077	966,196	5,941,537	6,907,733
		(+2.8%)	(-19.1%)	(-50.7%)	(-47.8%)	(+0.5%)	(-1.6%)	(-19.8%)	(-17.7%)
	$\rho = 0.0001$	2,376	18,367	55,483	73,850	41,231	968,295	2,774,310	3,742,605
		(+3.3%)	(-24.5%)	(-76.7%)	(-71.9%)	(+3.4%)	(-1.4%)	(-62.6%)	(-55.4%)
	$\rho = 0.001$	2,370	16,097	17,276	33,373	43,715	882,211	886,268	1,768,479
		(+3.0%)	(-33.9%)	(-92.8%)	(-87.3%)	(+9.7%)	(-10.2%)	(-88.0%)	(-78.9%)
	$\rho = 0.01$	2,524	14,041	1,303	15,344	54,713	722,116	294,544	1,016,660
		(+9.7%)	(-42.3%)	(-99.5%)	(-94.2%)	(+37.2%)	(-26.5%)	(-96.0%)	(-87.9%)
30 AP	(GDP)	2,915	34,577	603,468	638,045	42,237	1,044,339	8,614,036	9,658,375
	$\rho = 0.00001$	2,914	25,537	183,235	208,772	42,410	1,022,636	6,531,777	7,554,413
		(-0.0%)	(-26.1%)	(-69.6%)	(-67.3%)	(+0.4%)	(-2.1%)	(-24.2%)	(-21.8%)
	$\rho = 0.0001$	2,916	27,004	85,270	112,274	43,429	1,026,757	2,828,429	3,855,186
		(+0.0%)	(-21.9%)	(-85.9%)	(-82.4%)	(+2.8%)	(-1.7%)	(-67.2%)	(-60.1%)
	$\rho = 0.001$	2,987	24,322	27,173	51,495	46,256	940,132	867,430	1,807,562
		(+2.5%)	(-29.7%)	(-95.5%)	(-91.9%)	(+9.5%)	(-10.0%)	(-89.9%)	(-81.3%)
	$\rho = 0.01$	3,223	19,316	3,073	22,389	59,688	775,223	329,670	1,104,893
		(+10.5%)	(-44.1%)	(-99.5%)	(-96.5%)	(+41.3%)	(-25.8%)	(-96.2%)	(-88.6%)

(i)+(iii), whereas the "true" arrival delay of one-stop passengers ("total") is (i)+(ii)+(iii). Table 4 shows that the arrival delay is indeed much smaller than the costs of misconnections and that, as a result, the impact of (GDP-PAX) is similar on the simplified objective function and on the total arrival delay. This confirms that the insights derived from our results do not depend critically on this modeling simplification.

Table 4 One-stop passenger delay: objective (added connecting time) vs. total arrival delay ($w = 5$ hours).

Model	VMC					IMC				
	Miscon. (i)	Arr. delay (ii)	Sink (iii)	Obj. (i)+(iii)	Total (i)+(ii)+(iii)	Miscon. (i)	Arr. delay (ii)	Sink (iii)	Obj. (i)+(iii)	Total (i)+(ii)+(iii)
(GDP)	181,548	2,394	421,920	603,468	605,862	1,683,316	121,636	6,930,720	8,614,036	8,735,672
$\rho = 0.00001$	62,275	2,974	120,960	183,235	186,209	1,337,697	121,445	5,194,080	6,531,777	6,653,222
	(-65.7%)	(+24.2%)	(-71.3%)	(-69.6%)	(-69.3%)	(-20.5%)	(-0.2%)	(-25.1%)	(-24.2%)	(-23.8%)
$\rho = 0.0001$	44,950	2,347	40,320	85,270	87,617	888,749	99,075	1,939,680	2,828,429	2,927,504
	(-75.2%)	(-2.0%)	(-90.4%)	(-85.9%)	(-85.5%)	(-47.2%)	(-18.5%)	(-72.0%)	(-67.2%)	(-66.5%)
$\rho = 0.001$	19,973	2,447	7,200	27,173	29,620	311,590	108,057	555,840	867,430	975,487
	(-89.0%)	(+2.2%)	(-98.3%)	(-95.5%)	(-95.1%)	(-81.5%)	(-11.2%)	(-92.0%)	(-89.9%)	(-88.8%)
$\rho = 0.01$	3,073	4,151	0	3,073	7,224	94,950	227,463	234,720	329,670	557,133
	(-98.3%)	(+73.4%)	(-100%)	(-99.5%)	(-98.8%)	(-94.4%)	(+87.0%)	(-96.6%)	(-96.2%)	(-93.6%)

A potential concern about passenger-centric operations is that they could favor a subset of airlines (e.g., hub-and-spoke carriers) over others (e.g., point-to-point carriers). Figure 3 shows that this is not the case: the delay distribution across airlines is similar with $\rho = 0$ (i.e., under (GDP)) and $\rho > 0$ (i.e., under (GDP-PAX)) and each airline's shared of delayed flights also remains roughly identical (we partition airlines into 7 groups based on alliances). Whereas future research could capture equity considerations explicitly in passenger-centric operations (e.g., Barnhart et al. 2012, Bertsimas and Gupta 2016, Jacquillat and Vaze 2017), these results suggest at least that (GDP-PAX) does not induce more inequitable outcomes than (GDP).

(a) Total delay per airline   (b) Average delay per airline
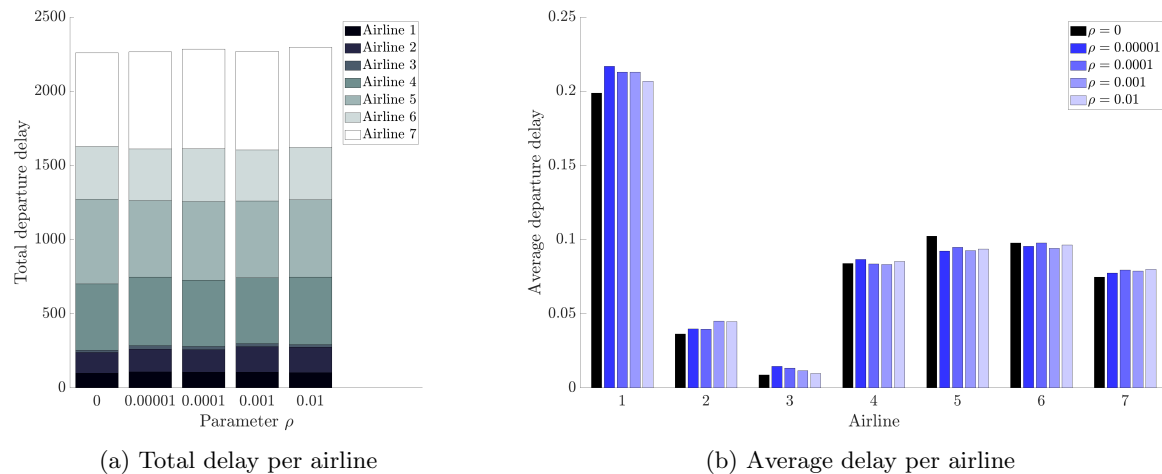
Figure 3   **Airline equity: delay distribution under (GDP) and (GDP-PAX)** ($|\mathcal{K}| = 30$, $w = 5$ **hours, VMC**).

Overall, these results show that integrating passenger-level considerations into existing ATFM technologies could alleviate the negative impact of air traffic congestion on passengers very significantly, at limited costs in terms of flight delays and inter-airline equity. By adjusting flight operating decisions based on passenger itineraries, one can avoid passenger misconnections and (especially under moderate congestion) absorb a higher share of delays on flights carrying fewer non-stop passengers. These benefits come at the cost of a (comparatively much smaller) increase in flight delay costs. Note, importantly, that airlines' costs depend both on flight delays (e.g., crew compensation, fuel burn) and passenger delays (e.g., passenger compensation). Ultimately, these results can thus support the determination of appropriate levels of flight and passenger delays—ideally in a collaborative decision-making environment involving the airlines and air traffic managers.

5.4. Flight-centric vs. Passenger-centric Decisions

We now characterize the optimal solution of (GDP-PAX) to discuss the drivers of passenger-centric GDP decisions. For any flight $i \in \mathcal{F}$, let d_i denote its scheduled departure time and let $\hat{d}_i$ (resp. $\hat{a}_i$) be its actual departure (resp. arrival) time. Also, let $\underline{d}_i$ denote its *earliest departure time* as the scheduled departure time d_i or the time when the aircraft is ready to operate flight i, whichever comes last:

$$\underline{d}_i = \begin{cases} \max\left(d_i, \hat{a}_j + \tau_{ji}\right) & \text{if there exists } j \in \mathcal{F} \text{ such that } (j,i) \in \mathcal{C} \\ d_i & \text{otherwise} \end{cases}$$

We define the *effective departure delay (EDD)* as the difference between actual and earliest departure times:

$$EDD_i = \hat{d}_i - \underline{d}_i, \quad \forall i \in \mathcal{F}$$

Inspired from the analytical insights from Section 2, we define the following three flight-level variables:

– *number of non-stop passengers*: The number of non-stop passengers in flight i is denoted by N_i and defined in Equation (34) ($\mathbb{1}$ denotes the indicator function). We posit that, as ρ increases, flights with more non-stop passengers will be subject to lower average EDD.

$$N_i = \sum_{p \in \mathcal{P}_N} \mathbb{1}\left[i = f_0(p)\right] n_p, \quad \forall i \in \mathcal{F} \tag{34}$$

– *number of active outgoing connections*: For any flight $i \in \mathcal{F}$, we define an outgoing connection as a passenger on a one-stop itinerary with flight i as its first leg. For example, all passengers booked on the LGA-ATL-LAX itinerary are counted as outgoing connections on the LGA-ATL flight. An "active" outgoing connection is defined such that the connection will be missed if flight i is subject to an EDD of 2 hours (8 periods) and the second flight is not delayed. The number of active outgoing connections of flight i is denoted by O_i and defined in Equation (35). We posit that, as ρ increases, flights with more active outgoing connections will be subject to lower average EDD.

$$O_i = \sum_{p \in \mathcal{P}_C} \mathbb{1}\left[i = f_1(p) \,\&\, \underline{d}_i + 8 > d_{f_2(p)}\right] n_p, \quad \forall i \in \mathcal{F} \tag{35}$$

– *number of active incoming connections*: For any flight $i \in \mathcal{F}$, we define an incoming connection as a passenger on a one-stop itinerary with flight i as its second leg. For example, all passengers booked on the LGA-ATL-LAX itinerary are counted as incoming connections on the ATL-LAX flight. An "active" incoming connection is defined such that the connection will be missed if flight i is operated at time $\underline{d}_i$, due to the delayed arrival of the first flight. The number of active incoming connections of flight i is denoted by I_i and defined in Equation (36). We posit that, as ρ increases, flights with more active incoming connections will be subject to higher average EDD.

$$I_i = \sum_{p \in \mathcal{P}_C} \mathbb{1}\left[i = f_2(p) \,\&\, \underline{d}_i < \hat{a}_{f_1(p)} + \sigma_{pi}\right] n_p, \quad \forall i \in \mathcal{F} \tag{36}$$

Figures 4, 5 and 6 plot the histograms of average EDD (in 15-minute periods) at the six busiest airports for different values of ρ, as a function of the number of non-stop passengers, outgoing connections and incoming connections, respectively. Absolute delays are highly sensitive to our (approximate) airport capacity estimates and should thus not be interpreted too literally. Instead, our insights focus on relative EDD variations.

The figures confirm our three hypotheses: all else equal, passenger-centric operations prioritize flights with more non-stop passengers, more outgoing connections and fewer incoming connections. Starting with non-stop passengers (Figure 4), the average EDD increases for flights with 50 non-stop passengers or less as ρ gets larger (by a factor of 2 at ATL, EWR and JFK). Vice versa, flights with most non-stop passengers have lower average delays. This is particularly striking at New York's airports (EWR, JFK, LGA), which face strong local demand and do not rely primarily on hub operations. Similarly, the average EDD of flights with no active outgoing connections increases with ρ, while that of flights with 40 or more active outgoing
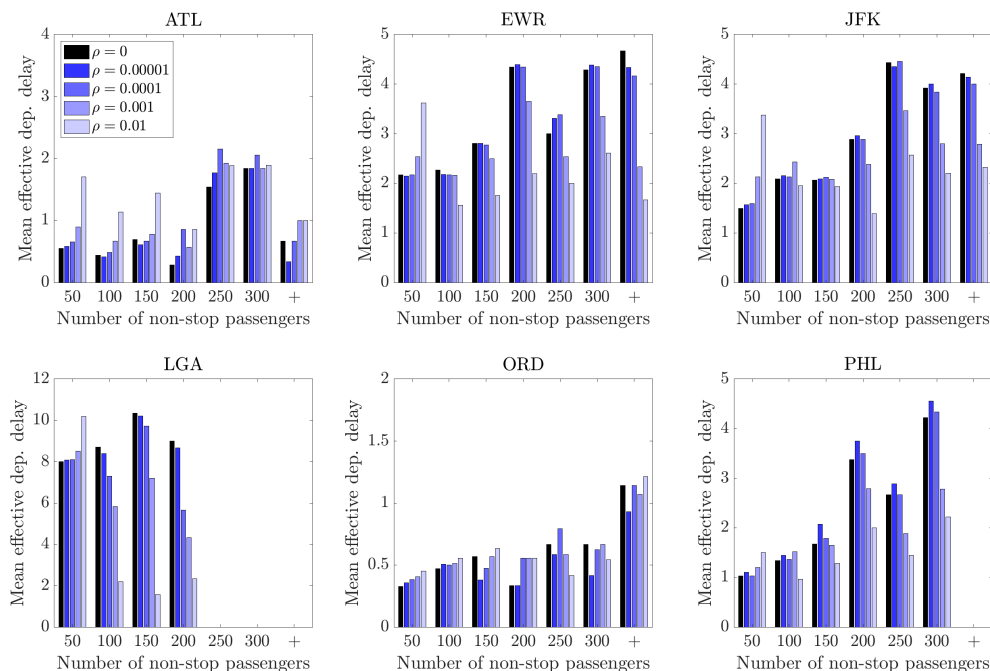
Figure 4   **Average EDD at six airports, as a function of non-stop passengers ($|\mathcal{K}| = 18$, $w = 5$ hours, IMC).**
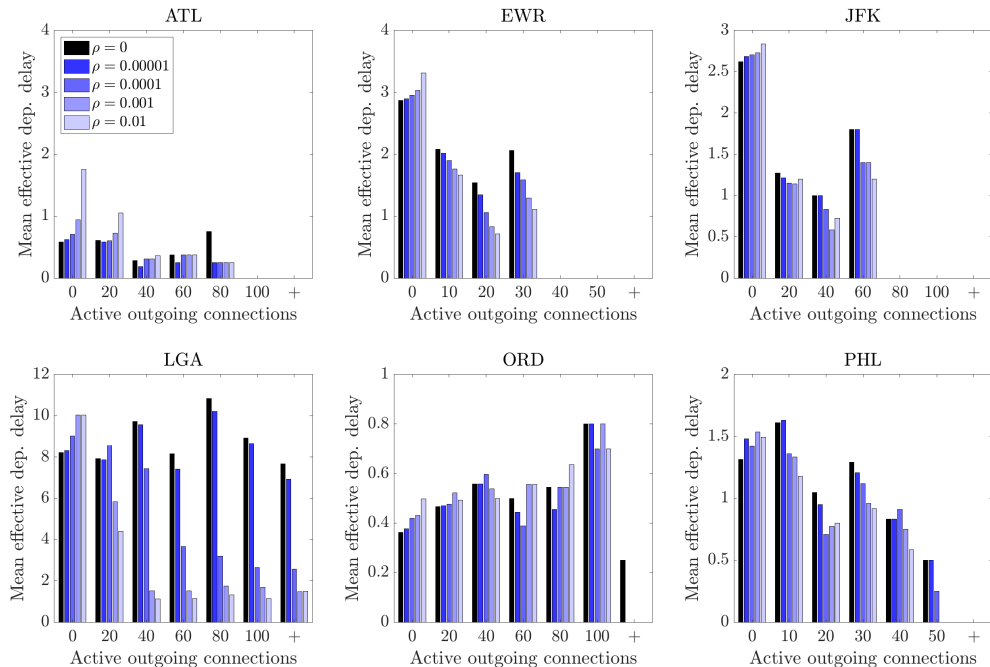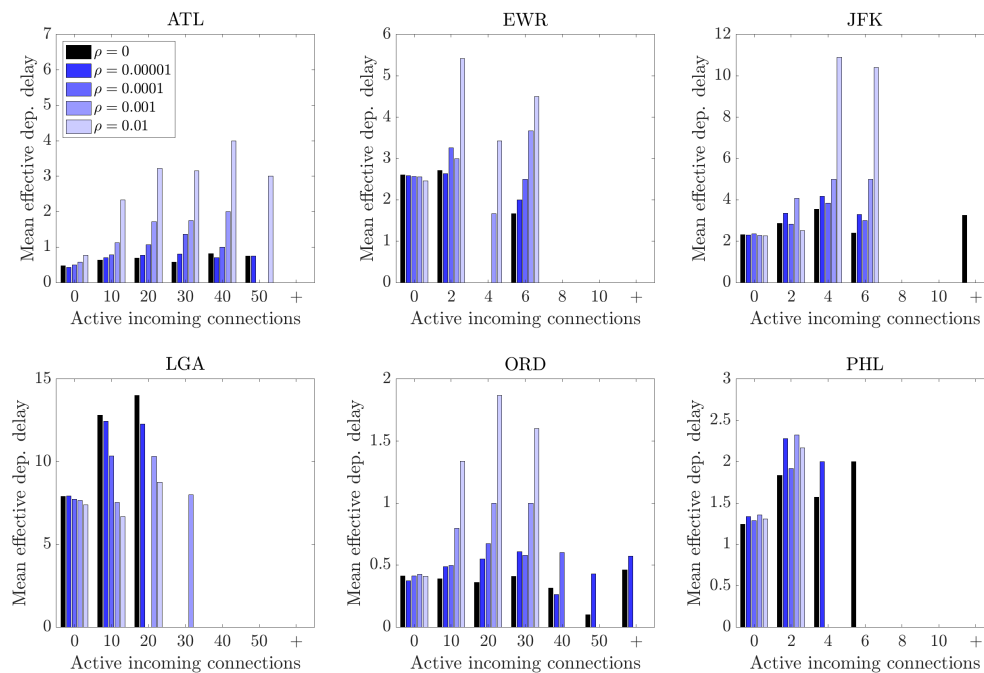

Figure 5   **Average EDD at six airports, as a function of outgoing connections ($|\mathcal{K}| = 18$, $w = 5$ hours, IMC).**

connections decreases with ρ (Figure 5). This effect is, by far, the strongest at LGA. Turning to Figure 6, we observe the opposite effect: as ρ increases, EDD increase, on average, for the flights with most active

Figure 6 **Average EDD at six airports, as a function of incoming connections** ($|\mathcal{K}| = 18$, $w = 5$ **hours, IMC**).

incoming connections. This effect is particularly strong at ATL and ORD, two major hubs of operations where passenger connectivity is critical. Note the one exception at LGA, where few passengers are connecting and passenger-centric operations are thus primarily driven by non-stop passengers and outgoing connections.

To synthesize these insights, Table 5 reports summary statistics across the full network. First, the number of passenger misconnections decreases dramatically as ρ gets larger. Even with the smallest values of ρ, (GDP-PAX) can cut misconnections by over 15% from the (GDP) baseline. As ρ increases, over 90% of system-wide misconnections are avoided. Next, as ρ gets larger, the mean EDD decreases for the flights with 100 or more non-stop passengers, and increases for the other flights. Similarly, it decreases for flights with active outgoing connections, but increases for the other flights. Last, and conversely, the mean EDD increases a lot for flights with active incoming connections, but remains essentially unchanged for the other flights.

These results confirm the analytical insights from Section 2 in large-scale networks of operations. Passenger-centric traffic flow management prioritize operations with many non-stop passengers, many outgoing connections, and few incoming connections. The other flights are de-prioritized, or even held deliberately on the ground if the benefits of avoiding misconnections outweigh the associated delay increases.

6. Integration of Predictive Analytics to Handle Data Unavailability

So far, we have assumed perfect knowledge of passenger itineraries. Ideally, such information could be provided by the airlines as part of Collaborative Decision Making (CDM) practices. But this also raises questions on the design of data-sharing mechanisms and airline participation. We thus investigate the robustness of the benefits of passenger-centric GDP when passenger itineraries are unknown. We thus *predict* passenger

Table 5   Summary statistics.

Metric	(GDP)	(GDP-PAX)			
		$\rho = 0.00001$	$\rho = 0.0001$	$\rho = 0.001$	$\rho = 0.01$
Missed passenger connections	9,511	8,066	4,603	2,044	670
Mean EDD, 100+ non-stop PAX	0.81 per.	0.81 per	0.81 per.	0.68 per.	0.53 per.
Mean EDD, ≤100 non-stop PAX	0.45 per.	0.45 per.	0.46 per.	0.50 per.	0.61 per.
Mean EDD, 1+ outgoing connections	0.47 per.	0.45 per.	0.38 per.	0.32 per.	0.29 per.
Mean EDD, no outgoing connections	0.52 per.	0.53 per.	0.57 per.	0.62 per.	0.75 per.
Mean EDD, 1+ incoming connections	0.66 per.	0.75 per.	0.90 per.	1.27 per.	2.50 per.
Mean EDD, no incoming connections	0.48 per.	0.47 per.	0.47 per.	0.47 per.	0.48 per.

itineraries based on historical data, and embed the resulting predictions into GDP optimization. Results suggest that this more realistic approach yields most of the benefits achieved under perfect information.

6.1.   Predictive Analytics for Passenger Itineraries

We propose analytics models to predict the number of passengers booked on each itinerary. To be conservative, we only leverage supply-side predictors (e.g., origin, destination, departure time, arrival time, etc.). With no doubt, the quality of the prediction could be enhanced with demand-side predictors (e.g., airfares, lag variables depicting historical passenger bookings). But, exactly like today's passenger itineraries are not necessarily known by air traffic managers, yesterday's demand-side features are not necessarily known either.

We separate the prediction task into two problems: (i) predicting the number of non-stop passengers on each flight and (ii) predicting the number of one-stop passengers on each flight pair. For both problems, the main challenge lies in feature engineering—building predictors that capture the patterns of passenger flows.

Features for predicting non-stop itineraries:

– *temporal information*: departure time, arrival time, scheduled block time;

– *spatial information*: origin airport (O), destination airport (D), $O-D$ distance;

– *airline information*: major airlines' names, each airline's traffic share in the network (measuring its size) and traffic shares at O and D (measuring its dominance at the origin and destination airports);

– *aircraft information*: number of seats in the aircraft;

– *traffic intensity*: traffic (number of flights over the previous week) at O, at D, and on the $O-D$ route; number of non-stop routes served by O and D; and

– *competition*: number of non-stop flights on the $O-D$ route during the day.

Features for predicting one-stop itineraries. We develop a dataset comprising all *potential* one-stop itineraries—all "compatible" flight pairs to/from the same connecting airport with a sufficient turnaround time. We eliminate all flight pairs flown by two airlines with very few passenger connections historically. To predict the number of passengers on each resulting one-stop itinerary, we consider the following features, where O, K and D refer to the origin, connection and destination airports:

– *temporal information*: departure time from O, departure time from K, arrival time at D;

– *spatial information*: $O-D$ distance, airport K, and indicator of whether K is a hub for the airline;

– *airline pair* (also indicative of the two flights are flown by the same airline or by partner airlines);

– *quality of connection*: layover duration, ratio between the connecting $O - K - D$ distance vs. the direct $O - D$ distance, ratio between the $O - K - D$ distance vs. the distance of the shortest available itinerary, and ratio between the $O - K - D$ distance vs. the distance of the shortest available connecting itinerary;

– *traffic intensity*: traffic at O and D; non-stop routes served by O and D; delays at O and D; and

– *competition*: number of non-stop flights and other one-stop options from O to D during the day.

Modeling. Armed with these features, we apply statistical learning methods to predict the number of passengers on each itinerary. We build a training set comprising all observations for the day immediately before the day under consideration—with 29,399 observations and 22 dependent variables for non-stop itineraries, and with 2,324,885 observations and 26 dependent variables for one-stop itineraries. We then evaluate the predictions on a test set comprising all observations during the day under consideration—with 29,061 observations for non-stop itineraries and 2,346,045 observations for one-stop itineraries. Given the problem size, we focus on simple—and scalable—methods: linear/logistic regression, ridge regression (Hoerl and Kennard 1970), Lasso (Tibshirani 1996), classification and regression trees (CART)(Breiman 2017), and random forests (Breiman 2001). We perform model selection via cross-validation (or out-of-bag estimation, for random forests), based on data from the previous day only but not from the day under consideration.

We first apply regression methods to predict the number of passengers on each non-stop itinerary.

We then strive to predict one-stop passenger itineraries on each flight pair. This problem, however, is more complex due to (i) the much larger number of observations (one per compatible flight pair vs. one per flight), (ii) imbalanced data (passengers connections occur on 5–7% of flight pairs), and (iii) noise in the data (depending on whether individuals vs. one family vs. several families are traveling, for instance). We first consider a one-step approach ("1S") for this problem, where we treat the problem as a regression problem. We then propose a two-step approach ("2S"), broken down into two-subproblems:

SP1. A classification subproblem ("2S–SP1") to predict the subset of all compatible flight pairs with at least one connection. Given the huge size of the training set (with over 2 million observations), we train the models on a stratified random sample of 10% of observations.

SP2. A conditional regression problem ("2S–SP2") to predict the number of one-stop passengers. We train the models on all flight pairs with at least one connecting passenger (156,845 training observations).
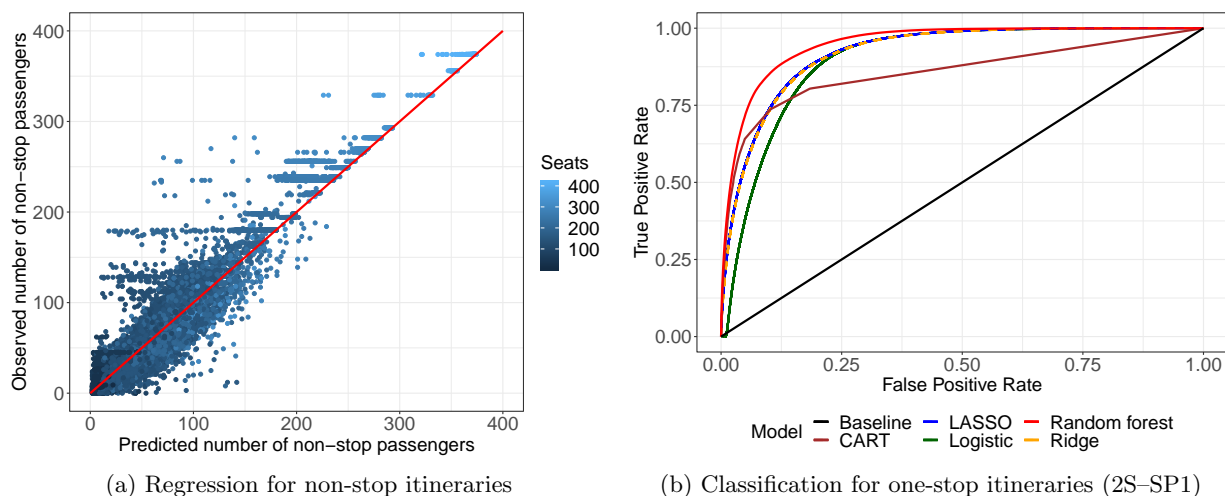
Results. Results are reported in Table 6. For non-stop itineraries, we obtain the best performance with the random forest. Figure 7a plots the predicted vs. actual values—most values fall near the 45-degree line. Ultimately, we get an out-of-sample R^2 of 0.945, a Mean Absolute Error (MAE) of 9 passengers and a Root Mean Squared Error (RMSE) of 15 passengers, suggesting excellent predictive performance.

For one-stop itineraries, the model predicts very well *which* flight pairs have connecting passengers but not the *number* of connecting passengers. Indeed, direct regression (1S approach) yields a relatively poor fit, with an R^2 of 0.15–0.30 with CART and random forests. The two-step (2S) approach, in contrast, shows that the classification outputs (SP1) are strong. Figure 7a plots the receiver operating characteristic (ROC) curves for each model. The out-of-sample area under the curve (AUC) reaches 95% for the random forest model. Stated differently, by leveraging only supply-side predictors and only 10% of observations from the

previous day, the model reliably distinguishes which flight pairs have connecting passengers and which ones do not. Then, the exact number of passengers remains more noisy, with an out-of-sample R^2 of only 0.25 (SP2). Yet, this remains a minor concern in view of our optimization problem because, first, the absolute error remains small (one passenger per itinerary on average) and, moreover, passenger-centric GDP mainly depend on the flight pairs with connecting passengers rather than the exact passenger count.

Table 6   **Cross-validated ("c.v.") and out-of-sample ("o.o.s.") performance of the predictive models.**

Problem	Metric	Evaluation	Linear model	Lasso	Ridge	CART	Random forests
Non-stop	R^2	(c.v.)	0.857	0.859	0.859	0.904	0.938
	R^2	(o.o.s.)	0.856	0.865	0.865	0.920	0.945
	MAE	(o.o.s.)	17.3	16.5	16.5	10.7	8.9
	RMSE	(o.o.s.)	24.3	23.6	23.6	18.2	15.1
One-stop (1S)	R^2	(c.v.)	0.143	0.144	0.144	0.300	0.237
	R^2	(o.o.s.)	-0.874	0.053	0.053	0.140	0.276
	MAE	(o.o.s.)	0.24	0.22	0.22	0.15	0.15
	RMSE	(o.o.s.)	0.83	0.58	0.58	0.56	0.51
One-stop (2S–SP1)	AUC	(c.v.)	0.953	0.953	0.953	0.862	0.955
	AUC	(o.o.s.)	0.898	0.922	0.920	0.858	0.950
One-stop (2S–SP2)	R^2	(c.v.)	0.203	0.205	0.205	0.244	0.166
	R^2	(o.o.s.)	0.153	0.146	0.146	0.152	0.254
	MAE	(o.o.s.)	1.0	1.0	1.0	1.0	0.9
	RMSE	(o.o.s.)	1.5	1.5	1.5	1.5	1.4

(a) Regression for non-stop itineraries

(b) Classification for one-stop itineraries (2S–SP1)

Figure 7   **Out-of-sample results for prediction of non-stop and one-stop passenger itineraries.**

We report the variable importance from our random forests in Figure 8. For non-stop itineraries, the main drivers of the predictions are the aircraft size, characteristics of the origin-destination pair, and the airline. For one-stop itineraries, the main drivers are the quality of the connection (layover duration and distance of connecting itinerary) and, to a lesser extent, characteristics of the origin-destination pair and the airlines.

Predictions. We now use the models to make test-set predictions. Note that the classification model SP1 does not return a single prediction but a probability of passenger connections on each flight pair. Accordingly, we build a range of predictions, corresponding to different true and false positive rates (Figure 7b).

(a) Regression for non-stop itineraries   (b) Classification for one-stop itineraries (2S–SP1)
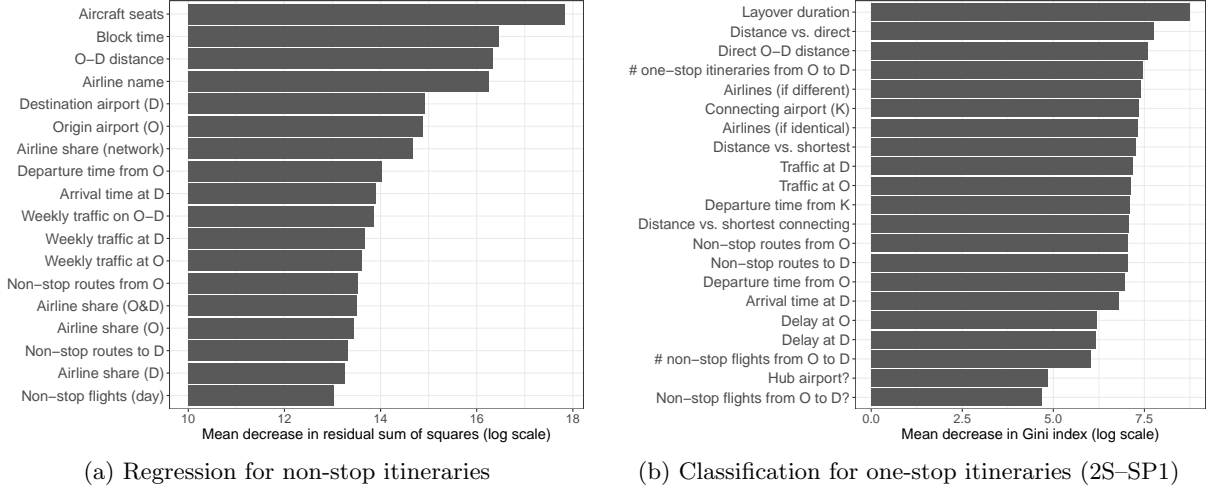
Figure 8   Variable importance for the random forest models.

One challenge remains: the predicted number of passengers on each flight can exceed the (known) capacity of the aircraft. We address these instances using two simple rules: for each flight with excess passengers, we first reduce the number of non-stop and one-stop passengers proportionally, and we then retain those one-stop passenger itineraries with the highest predicted probabilities. We detail this procedure in Algorithm 3.

Algorithm 3 Procedure to comply with aircraft capacities in passenger itinerary predictions.

- Define $\mathcal{P}_C^{\text{all}}$ as the set of all potential one-stop itineraries in the test set
- Define $q(i)$ as the non-stop itinerary coinciding with flight $i \in \mathcal{F}$
- Define $\mathcal{Q}_C(i) = \{p \in \mathcal{P} \mid f_1(p) = i \text{ or } f_2(p) = i\}$ as the subset of potential one-stop itineraries involving flight $i \in \mathcal{F}$
- Inputs: Probability $\widehat{prob}(p)$ that at least one passenger will connect on each potential one-stop itinerary $p \in \mathcal{P}_C^{\text{all}}$; number of predicted passengers $\hat{n}_p$ on each itinerary $p \in \mathcal{P}_N \cup \mathcal{P}_C^{\text{all}}$; aircraft capacity $\overline{\Omega}_i$ for each flight $i \in \mathcal{F}$
- On each flight with excess passengers, reduce the predicted number of non-stop passengers proportionally:

$$\hat{n}_{q(i)} \leftarrow \left\lfloor \frac{\hat{n}_{q(i)}}{\hat{n}_{q(i)} + \sum_{p \in \mathcal{Q}_C(i)} \hat{n}_p} \overline{\Omega}_i \right\rfloor, \ \forall i \in \mathcal{F} \text{ such that } \hat{n}_{q(i)} + \sum_{p \in \mathcal{Q}_C(i)} \hat{n}_p > \overline{\Omega}_i$$

while there exists $i \in \mathcal{F}$ such that $\hat{n}_{q(i)} + \sum_{p \in \mathcal{Q}_C(i)} \hat{n}_p > \overline{\Omega}_i$ **do**

- Find flight with the largest number of excess passengers: $i^* \leftarrow \arg\max \left\{ \hat{n}_{q(i)} + \sum_{p \in \mathcal{Q}_C(i)} \hat{n}_p - \overline{\Omega}_i \right\}$
- Sort $\mathcal{Q}_C(i)$ by non-increasing connection probability: $\mathcal{Q}_C(i) = \{p_1, \cdots, p_r\}$ so that $\widehat{prob}(p_1) \geq \cdots \geq \widehat{prob}(p_r)$
- Update the predicted number of one-stop itineraries accordingly from most likely to least likely itineraries:

$$\hat{n}_{p_k} \leftarrow \begin{cases} \hat{n}_{p_k} & \text{if } \sum_{l=1}^{k} \hat{n}_{p_l} \leq \overline{\Omega}_i - \hat{n}_{q(i)} \\ 0 & \text{otherwise} \end{cases}$$

end while

6.2.   GDP Optimization based on Predicted Passenger Itineraries

We now leverage our predicted non-stop and one-stop passenger itineraries to optimize GDP operations—using a similar rolling approach as in Algorithms 1 and 2. Specifically, we define two sets of passenger-level inputs: those based on actual itineraries ($\mathcal{P}_N$, $\mathcal{P}_C$, $f_1(p)$, $f_2(p)$, $\mathcal{R}_p$, n_p, δ_{pi}^C, δ_p^S and σ_{pi}, defined in Section 3),

and those based on predicted itineraries (denoted by $\mathcal{P}_N^{pr}$, $\mathcal{P}_C^{pr}$, $\hat{f}_1(p)$, $\hat{f}_2(p)$, $\mathcal{R}_p^{pr}$, $\hat{n}_p$ $\hat{\delta}_{pi}^C$, $\hat{\delta}_p^S$ and $\hat{\sigma}_{pi}$). At each iteration, we apply (GDP-PAX) to optimize flight operations based on *predicted* passenger itineraries. We then apply a passenger flow model to replicate passenger accommodations given the *actual* itineraries. Throughout, we track the final passenger accommodations, update the remaining number of passengers, and update the remaining aircraft capacities—using both predicted itineraries (for later optimization) and actual itineraries (for later passenger accommodations). That way, the optimization of flight-level operations remains blind to actual passenger itineraries throughout the algorithm—actual itineraries are solely used to track the impact of flight operations on passenger accommodations. We detail this procedure in Algorithm 4.

Algorithm 4 Rolling horizon algorithm to solve (GDP-PAX) with predicted passenger itineraries.

– Inputs: flight schedule, **actual** passenger itineraries ($\mathcal{P}_N$, $\mathcal{P}_C$, $f_1(p)$, $f_2(p)$, $\mathcal{R}_p$, n_p, δ_{pi}^C, δ_p^S and σ_{pi}), **predicted** passenger itineraries ($\mathcal{P}_N^{pr}$, $\mathcal{P}_C^{pr}$, $\hat{f}_1(p)$, $\hat{f}_2(p)$, $\mathcal{R}_p^{pr}$, $\hat{n}_p$ $\hat{\delta}_{pi}^C$, $\hat{\delta}_p^S$ and $\hat{\sigma}_{pi}$), aircraft capacities, and airport capacities

– Initialize final itineraries based on **actual** itineraries $\zeta_{pi} = 0$, $\forall p \in \mathcal{P}_C, i \in \mathcal{R}_p$, and $\zeta_p^S = 0$, $\forall p \in \mathcal{P}_C$

– Initialize final itineraries based on **predicted** itineraries $\hat{\zeta}_{pi} = 0$, $\forall p \in \mathcal{P}_C^{pr}, i \in \mathcal{R}_p^{pr}$, and $\hat{\zeta}_p^S = 0, \forall p \in \mathcal{P}_C^{pr}$

– Define rolling periods $u = 1, \cdots, U$ and duration of each look-ahead window w

    for $u = 1, \cdots, U$ **do**

      – Define flight sets: $\widetilde{\mathcal{F}}^a = \{i \in \mathcal{F} \,|\, u \le \tilde{a}_i \le u + w\}$ and $\widetilde{\mathcal{F}}^d = \left\{i \in \mathcal{F} \,|\, u \le \tilde{d}_i \le u + w\right\}$

      – Define passenger sets based on **predicted** itineraries: $\widetilde{\mathcal{P}}_C^{pr} = \left\{p \in \mathcal{P}_C^{pr} \,|\, \tilde{d}_{f_1(p)} \le u + w \text{ and } \sum_{i \in \mathcal{F}} \hat{\zeta}_{pi} + \hat{\zeta}_p^S < \hat{n}_p\right\}$, and subsets $\widetilde{\mathcal{P}}_C^{pr,A}$, $\widetilde{\mathcal{P}}_C^{pr,D}$, and $\widetilde{\mathcal{P}}_C^{pr,0}$ (Equations (24) to (26))

      – Solve (GDP-PAX) based on **predicted** itineraries: Equations (3)– (10), (27) (31), (13)– (18)

      – Update the arrival schedule $\tilde{a}_i$ for $i \in \widetilde{\mathcal{F}}^a$, and the departure schedule $\tilde{d}_i$ for $i \in \widetilde{\mathcal{F}}^d$

      – Fix final **predicted** itineraries; update $\hat{\zeta}_{pi}$ for $p \in \mathcal{P}_C^{pr}, i \in \mathcal{R}_p^{pr}$, and $\hat{\zeta}_p^S$ for $p \in \mathcal{P}_C^{pr}$ (Equations (19) and (20))

      – Update the number of remaining passengers and effective aircraft capacities based on **predicted** passenger itineraries (Equations (22) and (23))

      – Solve a passenger-flow model to reconstruct passenger itineraries, based on **actual** itineraries:

$$\min_{\lambda, z, z^S} \sum_{p \in \mathcal{P}_C} \left(\sum_{i \in \mathcal{R}_p} \delta_{pi}^C z_{pi} + \delta_p^S z_p^S \right) \quad \text{s.t.: Eq. (27)–(31), (13)–(16), (18)}$$

      – Fix final **actual** itineraries; update ζ_{pi} for $p \in \mathcal{P}_C, i \in \mathcal{R}_p$, and ζ_p^S for $p \in \mathcal{P}_C$ (Equations (19) and (20))

      – Update the number of remaining passengers and effective aircraft capacities based on **actual** passenger itineraries (Equations (22) and (23))

    end for

6.3. Experimental Results

We compare GDP performance when we ignore passenger itineraries (Algorithm 2), when we consider actual itineraries under perfect information (Algorithm 1), and when we consider predicted itineraries from historical data (Algorithm 4). We predict one-stop itineraries with the one-step (1S) and two-step (2S) approaches. For the 2S approach, we denote by κ the probability threshold above which we classify a flight pair as having at least one connecting passenger—a larger κ means fewer predicted connections and smaller true and false

Table 7   Summary statistics on predicted vs. actual passenger connections.

	Before Algorithm 3				After Algorithm 3			
Itineraries	Accuracy	TPR	FPR	$\|\mathcal{P}_C^{pr}\|$	Accuracy	TPR	FPR	$\|\mathcal{P}_C^{pr}\|$
Ignored	94.0%	0%	0%	0	94.0%	0%	0%	0
Perfect info.	100%	100%	0%	145,159	100%	100%	0%	145,159
Predicted (1S)	93.2%	62.8%	4.9%	195,458	93.4%	61.7%	4.6%	187,621
Predicted (2S, $\kappa = 0.1$)	85.8%	90.3%	14.5%	445,750	91.9%	75.9%	7.1%	262,486
Predicted (2S, $\kappa = 0.3$)	94.0%	69.0%	4.5%	194,933	94.3%	64.7%	3.8%	173,940
Predicted (2S, $\kappa = 0.5$)	95.2%	46.1%	1.7%	102,043	95.2%	45.3%	1.6%	99,577
Predicted (2S, $\kappa = 0.7$)	95.0%	23.5%	0.5%	43,294	95.0%	23.4%	0.5%	43,006
Predicted (2S, $\kappa = 0.9$)	94.2%	3.4%	0.03%	5,489	94.2%	3.4%	0.03%	5,489

Table 8   Flight delay costs and passenger delays with predicted vs. actual itineraries ($|\mathcal{K}| = 30$, $w = 5$ hours).

		VMC				IMC			
Instance	Itineraries	Flight		PAX		Flight		PAX	
$\rho = 0$	Ignored	2,915	(baseline)	638,045	(baseline)	42,237	(baseline)	9,658,375	(baseline)
$\rho = 0.00001$	Perfect info.	2,914	(-0.0%)	208,772	(-67.3%)	42,410	(+0.4%)	7,554,413	(-21.8%)
	Predicted (1S)	2,911	(-0.2%)	255,000	(-60.0%)	42,062	(-0.4%)	7,707,405	(-20.2%)
	Predicted (2S, $\kappa = 0.1$)	2,917	(+0.1%)	270,036	(-57.7%)	42,167	(-0.2%)	7,547,482	(-21.9%)
	Predicted (2S, $\kappa = 0.3$)	2,914	(-0.0%)	225,113	(-64.7%)	41,872	(-0.9%)	7,637,880	(-20.9%)
	Predicted (2S, $\kappa = 0.5$)	2,981	(+2.2%)	257,187	(-59.7%)	42,113	(-0.3%)	7,741,064	(-19.9%)
	Predicted (2S, $\kappa = 0.7$)	2,976	(+2.1%)	321,082	(-49.7%)	42,150	(-0.2%)	8,411,055	(-12.9%)
	Predicted (2S, $\kappa = 0.9$)	2,978	(+2.1%)	446,748	(-30.0%)	42,023	(-0.5%)	8,996,071	(-6.9%)
$\rho = 0.0001$	Perfect info.	2,916	(+0.0%)	112,274	(-82.4%)	43,429	(+2.8%)	3,855,186	(-60.1%)
	Predicted (1S)	2,928	(+0.4%)	168,865	(-73.5%)	43,719	(+3.5%)	4,592,322	(-52.5%)
	Predicted (2S, $\kappa = 0.1$)	2,865	(-1.7%)	136,028	(-78.7%)	44,354	(+5.0%)	3,998,088	(-58.6%)
	Predicted (2S, $\kappa = 0.3$)	2,893	(-0.8%)	173,305	(-72.8%)	43,866	(+3.9%)	4,122,429	(-57.3%)
	Predicted (2S, $\kappa = 0.5$)	2,967	(+1.8%)	202,342	(-68.3%)	43,270	(+2.4%)	4,777,335	(-50.5%)
	Predicted (2S, $\kappa = 0.7$)	2,944	(+1.0%)	277,570	(-56.5%)	42,671	(+1.0%)	6,195,492	(-35.9%)
	Predicted (2S, $\kappa = 0.9$)	2,944	(+1.0%)	519,157	(-18.6%)	42,294	(+0.1%)	8,754,018	(-9.4%)
$\rho = 0.001$	Perfect info.	2,987	(+2.5%)	51,495	(-91.9%)	46,256	(+9.5%)	1,807,562	(-81.3%)
	Predicted (1S)	3,025	(+3.8%)	124,513	(-80.5%)	46,511	(+10.1%)	2,996,205	(-69.0%)
	Predicted (2S, $\kappa = 0.1$)	3,082	(+5.7%)	97,624	(-84.7%)	48,162	(+14.0%)	2,492,435	(-74.2%)
	Predicted (2S, $\kappa = 0.3$)	2,994	(+2.7%)	116,631	(-81.7%)	46,683	(+10.5%)	2,821,288	(-70.8%)
	Predicted (2S, $\kappa = 0.5$)	3,052	(+4.7%)	143,548	(-77.5%)	45,519	(+7.8%)	3,074,194	(-68.2%)
	Predicted (2S, $\kappa = 0.7$)	3,026	(+3.8%)	242,746	(-62.0%)	44,559	(+5.5%)	4,338,981	(-55.1%)
	Predicted (2S, $\kappa = 0.9$)	2,993	(+2.7%)	539,179	(-15.5%)	43,005	(+1.8%)	7,851,957	(-18.7%)
$\rho = 0.01$	Perfect info.	3,223	(+10.5%)	22,389	(-96.5%)	59,688	(+41.3%)	1,104,893	(-88.6%)
	Predicted (1S)	3,334	(+14.4%)	100,973	(-84.2%)	57,712	(+36.6%)	1,938,715	(-79.9%)
	Predicted (2S, $\kappa = 0.1$)	3,476	(+19.3%)	69,031	(-89.2%)	65,355	(+54.7%)	1,797,646	(-81.4%)
	Predicted (2S, $\kappa = 0.3$)	3,268	(+12.1%)	88,116	(-86.2%)	59,646	(+41.2%)	1,940,665	(-79.9%)
	Predicted (2S, $\kappa = 0.5$)	3,166	(+8.6%)	97,790	(-84.7%)	55,025	(+30.3%)	2,478,973	(-74.3%)
	Predicted (2S, $\kappa = 0.7$)	3,083	(+5.8%)	210,171	(-67.1%)	52,630	(+24.6%)	3,903,246	(-59.6%)
	Predicted (2S, $\kappa = 0.9$)	3,107	(+6.6%)	484,536	(-24.1%)	50,081	(+18.6%)	7,014,553	(-27.4%)

positive rates (Figure 7b). Table 7 reports the accuracy, true positive rate, false positive rate, and number of connections with each predictive model, before and after correcting for aircraft capacities (Algorithm 3).

We now report the optimization results in Table 8 and in Figure 9. To ensure an apples-to-apples comparison, we report the *actual* passenger delays in all cases. The main observation is that the benefits of (GDP-PAX) are extremely robust to imperfect information on passenger itineraries. Indeed, the passenger delays obtained with predicted itineraries bridge most of the gap between those obtained with the baseline

(a) Flight delay costs (VMC)

(b) Passenger delays (VMC)

(c) Flight delay costs (IMC)

(d) Passenger delays (IMC)
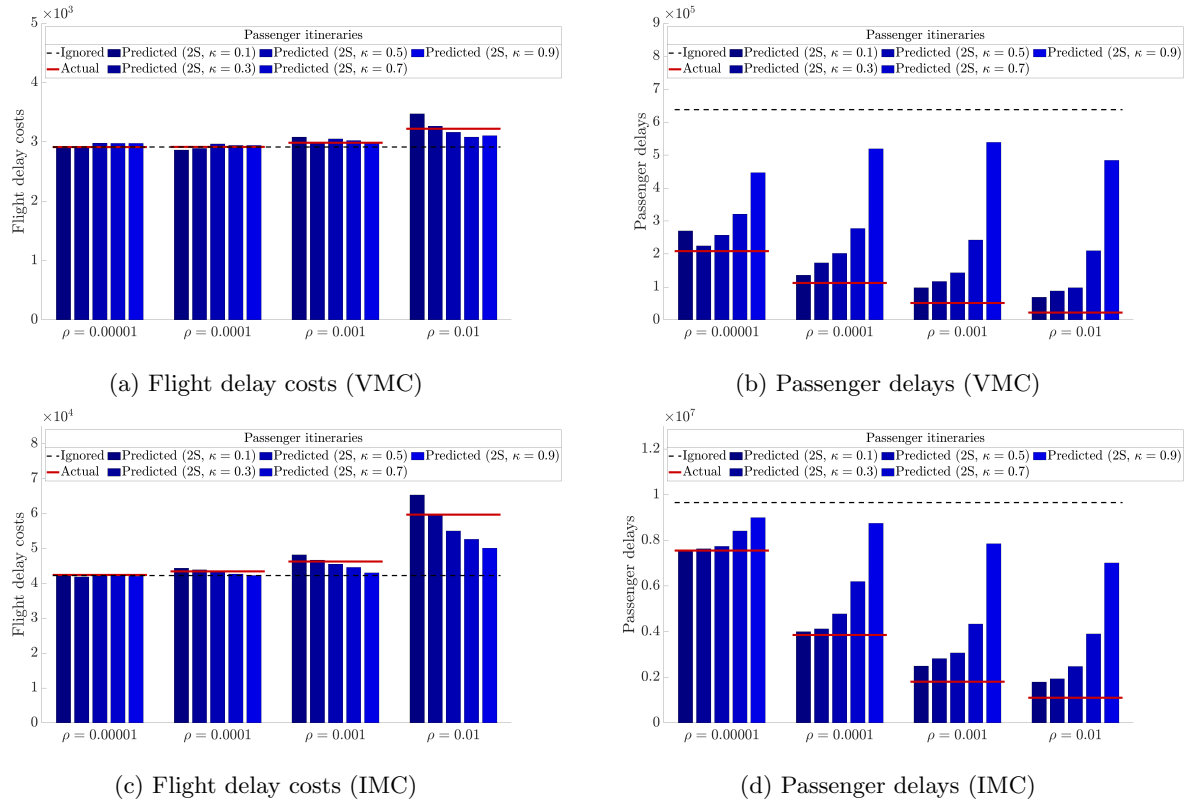
Figure 9 **Flight delay costs and passenger delays with predicted vs. actual itineraries ($|\mathcal{K}| = 30$, $w = 5$ hours).**

flight-centric model and those obtained with perfect knowledge of passenger itineraries. For instance, in VMC, passenger delays decrease by 55%–65% or by 70%–80% with predicted itineraries when they decrease by 67% and 82% with actual itineraries. In IMC, (GDP-PAX) yields virtually the same outcomes with predicted vs. actual itineraries, for smaller values of κ and ρ. Ultimately, our main observation holds: even under imperfect knowledge of passenger itineraries, passenger-centric operations can reduce passenger delays very significantly at moderate costs in terms of flight delays—by combining predictive analytics for reconstructing passenger flows and prescriptive analytics for GDP optimization.

Moreover, these benefits are very robust to the choice of the predictive model. Both the 1S and the 2S models achieve similar benefits. Yet, the 2S model generally results in lower aircraft delay and/or passenger delay, thus validating the practical benefits of the two-stage classification/regression approach considered in this paper. Moreover, the performance of the 2S approach depends on the probability threshold κ. Basically any value of $\kappa \leq 0.5$ reduces passenger delays significantly at moderate costs in terms of flight delays, as compared to the (GDP) baseline. The best results are obtained with small values of κ. With $\kappa = 0.9$, we consider a few connecting itineraries in the set $\mathcal{P}_C^{pr}$, so the outcome is closer to the baseline flight-centric (GDP) model. When $\kappa \leq 0.5$, the (GDP-PAX) model anticipates more one-stop itineraries. The prediction accuracy does not necessarily increase (e.g., it is lowest with $\kappa = 0.1$) but passenger delays can decrease very significantly (at the cost of slightly higher flight delays). Ultimately, the benefits of preserving passenger connectivity (by predicting more connections) can outweigh the costs of prediction errors.

Finally, the value of information on passenger itineraries increases with ρ. Consider for instance the results with $\kappa = 0.3$, which lead in all cases to almost-identical flight delays with predicted vs. actual itineraries. With predicted itineraries, passenger delays are higher than with actual itineraries by 8% when $\rho = 0.00001$, by 54% when $\rho = 0.0001$, by 126% when $\rho = 0.001$, and by 294% when $\rho = 0.01$, in VMC (in IMC, the corresponding numbers are 1.1%, 7%, 56%, and 76%). In other words, our predict-then-optimize approach captures the first-order effects of passenger-centric operations. For instance, if one flight has incoming connections from dozens of incoming flights and another one has only one incoming connection, even imperfect predictions will likely prioritize the departure of the latter flight. But when one attempts to avoid nearly all passenger misconnections, the knowledge of actual itineraries becomes much more critical—there is only so much one can do with predicted itineraries. For instance, if one flight has one incoming connection and another flight has none, the operating decision becomes much more sensitive to prediction quality.

From a practical standpoint, these results suggest that the proposed passenger-centric approach to GDP operations can provide strong benefits *even in the current operating environment* where passenger itineraries are not necessarily known by air traffic managers. The combination of predictive and prescriptive analytics can still make operations more consistent with passenger preferences—by leveraging historical data on passenger flows, statistical learning methods, and a passenger-centric optimization approach. Yet, additional benefits could be derived from the knowledge of actual itineraries, highlighting the value of data sharing between airlines and air traffic managers—for instance through the Collaborative Decision Making paradigm.

7. Conclusion

Vehicle-centric operations may not lead to optimal outcomes from the perspective of passengers, especially with connecting itineraries. To address this challenge, this paper balances vehicle-level and passenger-level objectives in traffic flow management, by combining predictive and prescriptive analytics. We formulated a large-scale integer optimization model that captures passenger accommodations in ground delay programs (GDP). We also addressed issues of data unavailability by predicting passenger itineraries using statistical learning methods and embedding the predictions into GDP optimization. Results suggest that passenger-level considerations can be incorporated into GDP optimization at limited computational costs, yielding very large reductions in passenger delays at comparatively small increases in flight delay costs. These benefits are driven by prioritizing flights carrying more non-stop passengers, more outgoing connections and fewer incoming connections—resulting in a decrease in non-stop passenger delay and a sharp reduction in passenger misconnections. Moreover, these results are robust to imperfect information—most of the benefits of passenger-centric operations hold with predicted (as opposed to actual) passenger itineraries.

These results suggest GDP initiatives can be enhanced by explicitly accounting for information on passenger itineraries—with significant benefits to airlines, passengers and other stakeholders. Even in the current operating environment where passenger itineraries are privately known by the airlines, these benefits can be captured by leveraging historical data, state-of-the-art statistical learning methods, and a passenger-centric optimization approach such as the one proposed in this paper. Preferably, passenger itineraries could be added to the data exchange underlying the existing Collaborative Decision Making paradigm to make resulting operations more consistent with air traffic managers', airlines' and passengers' preferences.

Finally, the limitations of this motivate further research on passenger-centric transportation. First, this paper has focused on the deterministic multi-airport ground-holding problem. The passenger accommodation model could be incorporated into more complex air traffic flow management (ATFM) models that capture en-route routing and operating uncertainty. Second, this paper has shown the benefits of data sharing on passenger itineraries between the airlines and air traffic managers. An open question involves designing efficient and equitable mechanisms to support such data sharing. Third, this paper has showed that passenger connectivity could be largely maintained through minor adjustments in ATFM interventions. In practice, however, many operating decisions are decentralized to the airlines, which can swap and cancel flights following ATFM decisions—based on flight-level as well as passenger-level information. The purpose of this research is not to fundamentally alter this well-accepted operating paradigm, but to show the benefits of accounting for passenger-level information at the upstream level. One might expect that these benefits would carry over at the downstream level, once the airlines re-optimize their own networks of flights as part of their disruption management processes. Yet, this question opens important avenues for future research to combine the proposed passenger-centric ATFM approach with downstream airline disruption recovery processes. This paper provides the methodological foundation to address such problems—with the ultimate objectives of mitigating delay costs borne by operators, providers and passengers across transportation systems.

References

Aouad A, Farias VF, Levi R, 2015 *Assortment optimization under consider-then-choose choice models. Available at SSRN 2618823* .

Balakrishnan H, Chandran B, 2018 *Optimal Large-Scale Air Traffic Flow Management. Working Paper* .

Ball M, Barnhart C, Dresner M, Hansen M, Neels K, Odoni A, Peterson E, Sherry L, Trani A, Zou B, 2010 *Total Delay Impact Study.* Technical report, National Center of Excellence for Aviation Operations Research, College Park, MD.

Ball M, Barnhart C, Nemhauser G, Odoni A, 2007 *Air Transportation: Irregular Operations and Control.* Barnhart C, Laporte G, eds., *Handbook in Operations Research & Management Science*, volume 14, 1–67 (Elsevier).

Ball M, Hoffman R, Odoni A, Rifkin R, 2003 *A Stochastic Integer Program with Dual Network Structure and its Application to the Ground-Holding Problem. Operations Research* 51(1):167–171.

Ball M, Swaroop P, Barnhart C, Yan C, Hansen M, Kang L, Liu Y, Vaze V, 2015 *Service level expectation setting for air traffic flow management: Practical challenges and benefits assessment. Proceedings of the 12th USA/Europe air traffic management R&D seminar.*

Barnhart C, Bertsimas D, Caramanis C, Fearing D, 2012 *Equitable and Efficient Coordination in Traffic Flow Management. Transportation Science* 46(2):262–280.

Barnhart C, Fearing D, Vaze V, 2014 *Modeling Passenger Travel and Delays in the National Air Transportation System. Operations Research* 62(3):580–601.

Bertsekas D, 2012 *Dynamic Programming and Optimal Control*, volume II (Athena Scientific), 4th edition.

Bertsimas D, Delarue A, Martin S, 2019 *Optimizing schools' start time and bus routes. Proceedings of the National Academy of Sciences* 116(13):5943–5948.

Bertsimas D, Farias V, Trichakis N, 2012 *On the Efficiency-Fairness Trade-Off. Management Science* 58(12):2234–2250.

Bertsimas D, Gupta S, 2016 *Fairness and Collaboration in Network Air Traffic Flow Management: An Optimization Approach. Transportation Science* 50(1):57–76.

Bertsimas D, Kallus N, 2019 *From predictive to prescriptive analytics. Management Science* .

Bertsimas D, Lulli G, Odoni A, 2011 *An Integer Optimization Approach to Large-Scale Air Traffic Flow Management. Operations Research* 59(1):211–227.

Bertsimas D, Stock Patterson S, 1998 *The Air Traffic Flow Management Problem with Enroute Capacities. Operations Research* 46(3):406–422.

Besbes O, Gur Y, Zeevi A, 2016 *Optimization in online content recommendation services: Beyond click-through rates. Manufacturing & Service Operations Management* 18(1):15–33.

Bratu S, Barnhart C, 2006 *Flight Operations Recovery: New Approaches Considering Passenger Recovery. Journal of Scheduling* 9(3):279–298.

Breiman L, 2001 *Random forests. Machine learning* 45(1):5–32.

Breiman L, 2017 *Classification and regression trees* (Routledge).

Dahan M, Saurabh A, Barnhart C, Justice J, Lee A, 2020 *Uncertainty-Aware Routing of Aerial Sensors for Infrastructure Damage Inspection. Working paper* .

Elmachtoub AN, Grigas P, 2017 *Smart "predict, then optimize". arXiv preprint arXiv:1710.08005* .

Federal Aviation Administration, 2004 *Airport Capacity Benchmark Report.* Technical report.

Federal Aviation Administration, 2013 *Aviation System Performance Metrics (ASPM) database. Accessed April 4, 2013.* Available at: https://aspm.faa.gov/apm/sys/main.asp.

Ferreira KJ, Lee BHA, Simchi-Levi D, 2016 *Analytics for an online retailer: Demand forecasting and price optimization. Manufacturing & Service Operations Management* 18(1):69–88.

Gallien J, Mersereau AJ, Garro A, Mora AD, Vidal MN, 2015 *Initial shipment decisions for new products at Zara. Operations Research* 63(2):269–286.

Gilbo E, 1993 *Airport Capacity: Representation, Estimation, Optimization. IEEE Transactions on Control Systems Technology* 1(3):144–154.

Hoerl AE, Kennard RW, 1970 *Ridge regression: Biased estimation for nonorthogonal problems. Technometrics* 12(1):55–67.

Jacquillat A, Odoni A, 2015 *An Integrated Scheduling and Operations Approach to Airport Congestion Mitigation. Operations Research* 63(6):1390–1410.

Jacquillat A, Odoni A, Webster M, 2017 *Dynamic Control of Runway Configurations and of Arrival and Departure Service Rates at JFK Airport under Stochastic Queue Conditions. Transportation Science* 51(1):155–176.

Jacquillat A, Vaze V, 2017 *Inter-airline Equity in Airport Scheduling Interventions. Transportation Science* Articles in advance.

Jones J, Lovell D, Ball M, 2017 *Stochastic Optimization Models for Transferring Delay Along Flight Trajectories to Reduce Fuel Usage. Transportation Science* Articles in advance.

Liu S, He L, Shen ZJM, 2020 *On-time last mile delivery: Order assignment with travel time predictors. Management Science* in press.

Lulli G, Odoni A, 2007 *The European Air Traffic Flow Management Problem. Transportation Science* 41(4):431–443.

Marla L, Vaaben B, Barnhart C, 2017 *Integrated Disruption Management and Flight Planning to Trade Off Delays and Fuel Burn. Transportation Science* Articles in advance.

Mukherjee A, Hansen M, 2007 *A Dynamic Stochastic Model for the Single Airport Ground Holding Problem. Transportation Science* 41(4):444–456.

Odoni A, 1987 *The Flow Management Problem in Air Traffic Control.* Odoni A, Bianco L, Szego G, eds., *Flow Control of Congested Networks* (Springer-Verlag).

Pellegrini P, Castelli L, Pesenti R, 2012 *Secondary Trading of Airport Slots as a Combinatorial Exchange. Transportation Research Part E: Logistics and Transportation Review* 48(5):1009–1022.

Pyrgiotis N, 2011 *A Stochastic and Dynamic Model of Delay Propagation Within an Airport Network For Policy Analysis.* Ph.D. thesis, Massachusetts Institute of Technology.

Richetta O, Odoni A, 1993 *Solving Optimally the Static Ground-Holding Policy Problem in Air Traffic Control. Transportation Science* 27(3):228–237.

Simaiakis I, 2012 *Analysis, Modeling and Control of the Airport Departure Process.* Ph.D. thesis, Massachusetts Institute of Technology.

Swaroop P, Zou B, Ball M, Hansen M, 2012 *Do More US Airports Need Slot Controls? A Welfare Based Approach to Determine Slot Levels. Transportation Research Part B: Methodological* 46(9):1239–1259.

Terrab M, Odoni A, 1993 *Strategic Flow Management for Air Traffic Control. Operations Research* 41(1):138–152.

Tibshirani R, 1996 *Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological)* 58(1):267–288.

Vaze V, Barnhart C, 2012 *Modeling Airline Frequency Competition for Airport Congestion Mitigation. Transportation Science* 46(4):512–535.

Vossen T, Ball M, 2006 *Slot Trading Opportunities in Collaborative Ground Delay Programs. Transportation Science* 40(1):29–43.

Vossen T, Hoffman R, Mukherjee A, 2012 *Air Traffic Flow Management. Quantitative Problem Solving Methods in the Airline Industry*, volume 169 of *International Series in Operations Research & Management Science*, 385–453 (Springer US).

Vranas P, Bertsimas D, Odoni A, 1994 *The Multi-airport Ground-holding Problem in Air Traffic Control. Operations Research* 42(2):249–261.

Wambsganss M, 1996 *Collaborative Decision Making through Dynamic Information Transfer. Air Traffic Control Quarterly* 4(2):107–123.

Appendix A: Proof of Statements from Section 2

A.1. Proof of Lemma 1

Let t_s be the time of the sth transition, for $s = 0, \cdots, n+m$ (by convention, $t_0 = 0$). Let $(\tilde{x}^s, \tilde{y}^s)$ be the state at time t_s. Finally, $\mathbb{1}^{(i)}_{(\tilde{x}^s, \tilde{y}^s)}$ takes the value 1 if the next event is the arrival of vehicle i.

The duration between decision epochs is exponentially distributed, with rate $\sum_{i \in \mathcal{A}(x)} \lambda_i$ if $\pi(x,y) = 0$, and rate $\mu + \sum_{i \in \mathcal{A}(x)} \lambda_i$ if $\pi(x,y) \in \mathcal{B}(y)$. The continuous-time cost function given below. The first term reflects the wait times accrued between t_s and t_{s+1} by the $\tilde{y}^s_j$ passengers waiting for vehicle j to depart, for each $j \in \mathcal{B}(\tilde{y}^s)$. The second term captures misconnection costs: if the next event is the arrival of vehicle $i \in \mathcal{A}(\tilde{x}^s)$, then, for each already-departed vehicle $j \notin \mathcal{B}(\tilde{y}^s)$, γ_{ij} passengers miss their connection at cost c_j.

$$\min \quad \mathbb{E}\left\{ \sum_{s=0}^{n+m} \left[\int_{t_s}^{t_{s+1}} e^{-\beta t} \times \left(\sum_{j \in \mathcal{B}(\tilde{y}^s)} \tilde{y}^s_j \right) + \sum_{i \in \mathcal{A}(\tilde{x}^s)} \mathbb{1}^{(i)}_{(\tilde{x}^s, \tilde{y}^s)} \times \sum_{j \notin \mathcal{B}(\tilde{y}^s)} \gamma_{ij} c_j \right] \right\} \tag{37}$$

We follow the procedure outlined by Bertsekas (2012) for continuous-time Markov Decision Processes.

We decompose the time interval $[0, t_s]$ by writing $[0, t_s] = \cup_{q=1}^{s} [t_{q-1}, t_q]$. We first express $E(e^{-\beta t_s})$, as follows. We use the facts that the transition rate from t_{q-1} to t_q is equal to $\mu + \sum_{i=1}^{n} \lambda_i$, that $E(e^{-\beta \tau}) = \frac{\xi}{\beta + \xi}$ if τ follows an exponential distribution with parameter ξ, and that all transition processes are independent.

$$E\left(e^{-\beta t_s}\right) = E\left[e^{-\beta \sum_{q=1}^{s}\left(t_q - t_{q-1}\right)}\right] = \prod_{q=1}^{s} e^{-\beta \left(t_q - t_{q-1}\right)} = \left(\frac{\mu + \sum_{i=1}^{n} \lambda_i}{\beta + \mu + \sum_{i=1}^{n} \lambda_i} \right)^s$$

We now use this result to derive $E\left(\int_{t_s}^{t_{s+1}} e^{-\beta t} dt \right)$:

$$E\left(\int_{t_s}^{t_{s+1}} e^{-\beta t} dt \right) = \frac{E\left(e^{-\beta t_s}\right) - E\left(e^{-\beta(t_{s+1})}\right)}{\beta} = E\left(e^{-\beta t_s}\right) \frac{1 - E\left(e^{-\beta(t_{s+1} - t_s)}\right)}{\beta}$$

$$= \left(\frac{\mu + \sum_{i=1}^{n} \lambda_i}{\beta + \mu + \sum_{i=1}^{n} \lambda_i} \right)^s \frac{1 - \frac{\mu + \sum_{i=1}^{n} \lambda_i}{\beta + \mu + \sum_{i=1}^{n} \lambda_i}}{\beta} = \left(\frac{\mu + \sum_{i=1}^{n} \lambda_i}{\beta + \mu + \sum_{i=1}^{n} \lambda_i} \right)^s \frac{1}{\beta + \mu + \sum_{i=1}^{n} \lambda_i}$$

Finally, $E\left(\mathbb{1}^{(i)}_{(\tilde{x}^s, \tilde{y}^s)} \right)$ is equal to the transition probability from (x,y) to $f_i(x,y)$, i.e., $\frac{\lambda_i}{\mu + \sum_{k=1}^{n} \lambda_k}$. Therefore:

$$z = \sum_{s=0}^{n+m} \left[\left(\frac{\mu + \sum_{i=1}^{n} \lambda_i}{\beta + \mu + \sum_{i=1}^{n} \lambda_i} \right)^s \frac{\sum_{j \in \mathcal{B}(\tilde{y}^s)} \tilde{y}^s_j}{\beta + \mu + \sum_{i=1}^{n} \lambda_i} + \left(\frac{\mu + \sum_{i=1}^{n} \lambda_i}{\beta + \mu + \sum_{i=1}^{n} \lambda_i} \right)^{s+1} \sum_{i \in \mathcal{A}(\tilde{x}^s)} \frac{\lambda_i}{\mu + \sum_{k=1}^{n} \lambda_k} \left(\sum_{j \notin \mathcal{B}(\tilde{y}^s)} \gamma_{ij} c_j \right) \right]$$

$$z = \frac{1}{\beta + \mu + \sum_{i=1}^{n} \lambda_i} \sum_{s=0}^{n+m} \left(\frac{\mu + \sum_{i=1}^{n} \lambda_i}{\beta + \mu + \sum_{i=1}^{n} \lambda_i} \right)^s \left[\sum_{j \in \mathcal{B}(\tilde{y}^s)} \tilde{y}^s_j + \sum_{i \in \mathcal{A}(\tilde{x}^s)} \lambda_i \left(\sum_{j \notin \mathcal{B}(\tilde{y}^s)} \gamma_{ij} c_j \right) \right]$$

We now use this formula to derive the Bellman equation given the transitions of the system given by:

$$\text{if } \pi(x,y) = j \in \mathcal{B}(y) \quad \text{then:} \quad \begin{cases} p_{(x,y) \to (x,y)} &= \frac{\sum_{k \notin \mathcal{A}(x)} \lambda_k}{\mu + \sum_{k=1}^{n} \lambda_k} \quad \forall i \in \mathcal{A}(x) \\ p_{(x,y) \to f_i(x,y)} &= \frac{\lambda_i}{\mu + \sum_{k=1}^{n} \lambda_k} \\ p_{(x,y) \to g_j(x,y)} &= \frac{\mu}{\mu + \sum_{k=1}^{n} \lambda_k} \end{cases}$$

$$\text{if } \pi(x,y) = 0 \quad \text{then:} \quad \begin{cases} p_{(x,y) \to (x,y)} &= \frac{\mu + \sum_{k \notin \mathcal{A}(x)} \lambda_k}{\mu + \sum_{k=1}^{n} \lambda_k} \quad \forall i \in \mathcal{A}(x) \\ p_{(x,y) \to f_i(x,y)} &= \frac{\lambda_i}{\mu + \sum_{k=1}^{n} \lambda_k} \end{cases}$$

For notational ease, we denote by $\tilde{J}_0(x,y)$ and $\tilde{J}_l(x,y)$ the future cost-to-go (from the next transition onward) assuming that the current decision is 0 and $l \in \mathcal{B}(y)$, respectively. We have:

$$\tilde{J}_0(x,y) = \frac{\mu + \sum_{k \notin \mathcal{A}(x)} \lambda_k}{\mu + \sum_{k=1}^{n} \lambda_k} J^*(x,y) + \sum_{i \in \mathcal{A}(x)} \frac{\lambda_i}{\mu + \sum_{k=1}^{n} \lambda_k} J^*(f_i(x,y))$$

$$\tilde{J}_l(x,y) = \frac{\sum_{k\notin\mathcal{A}(x)}\lambda_k}{\mu+\sum_{k=1}^n\lambda_k}J^*(x,y) + \sum_{i\in\mathcal{A}(x)}\frac{\lambda_i}{\mu+\sum_{k=1}^n\lambda_k}J^*(f_i(x,y)) + \frac{\mu}{\mu+\sum_{k=1}^n\lambda_k}J^*(g_j(x,y)) \qquad \forall l\in\mathcal{B}(y)$$

$$J^*(x,y) = \frac{1}{\beta+\mu+\sum_{i=1}^n\lambda_i}\left[\sum_{j\in\mathcal{B}(y)}y_j + \sum_{i\in\mathcal{A}(x)}\lambda_i\left(\sum_{j\notin\mathcal{B}(y)}\gamma_{ij}c_j\right)\right] + \frac{\mu+\sum_{i=1}^n\lambda_i}{\beta+\mu+\sum_{i=1}^n\lambda_i}\min\left\{\tilde{J}_0(x,y)\min_{l\in\mathcal{B}(y)}\left(\tilde{J}_l(x,y)\right)\right\}$$

$$J^*(x,y) = \frac{1}{\beta+\mu+\sum_{i=1}^n\lambda_i}\left[\sum_{j\in\mathcal{B}(y)}y_j + \sum_{i\in\mathcal{A}(x)}\lambda_i\left(\sum_{j\notin\mathcal{B}(y)}\gamma_{ij}c_j\right) + \sum_{i\notin\mathcal{A}(x)}\lambda_iJ^*(x,y) + \sum_{i\in\mathcal{A}(x)}\lambda_iJ^*(f_i(x,y))\right.$$
$$\left. + \mu\min\left\{J^*(x,y), \min_{l\in\mathcal{B}(y)}\left(J^*(g_l(x,y))\right)\right\}\right]$$

After eliminating the $J^*(x,y)$ terms on both sides of the equation above, we obtain:

$$J^*(x,y) = \frac{1}{\beta+\mu+\sum_{i\in\mathcal{A}(x)}\lambda_i}\left[\sum_{j\in\mathcal{B}(y)}y_j + \sum_{i\in\mathcal{A}(x)}\lambda_i\left(\sum_{j\notin\mathcal{B}(y)}\gamma_{ij}c_j\right) + \sum_{i\in\mathcal{A}(x)}\lambda_iJ^*(f_i(x,y))\right.$$
$$\left. \mu\min\left\{J^*(x,y), \min_{l\in\mathcal{B}(y)}\left(J^*(g_l(x,y))\right)\right\}\right]$$

This completes the proof. $\square$

In the remainder of this appendix, we use the notation of $\tilde{J}_0(x,y)$ and $\tilde{J}_l(x,y)$ to refer to the future cost-to-go assuming that the current decision is 0 and $l\in\mathcal{B}(y)$, respectively.

A.2. Proof of Proposition 1

For notational ease, we replace λ_1, c_1, y_1 and γ_{11} by λ, c, y and γ, respectively, By definition:

$$J^*(0,T) = 0$$
$$J^*(1,T) = \frac{1}{\beta+\lambda}\left[\lambda J^*(0,T) + \lambda\gamma c\right] = \frac{\lambda}{\beta+\lambda}\gamma c$$
$$J^*(0,y+\gamma) = \frac{1}{\beta+\mu}\left[(y+\gamma) + \mu\min\{J^*(0,y+\gamma), J^*(0,T)\}\right] = \frac{y+\gamma}{\beta+\mu}$$
$$J^*(1,y) = \frac{1}{\beta+\mu+\lambda}\left[y + \lambda J^*(0,y+\gamma) + \mu\min\{J^*(1,y), J^*(1,T)\}\right]$$

We obtain that $J^*(0,y+\gamma) \geq J^*(0,T) = 0$ for all $y\geq 0$, so the optimal policy satisfies $\pi^*(0,y+\gamma) = 1$ for all $y\geq 0$. Moreover, we have: $J^*(0,y+\gamma) = \frac{y+\gamma}{\beta+\mu}$. This simply states that, once the incoming vehicle has arrived, the optimal policy consists of immediately operating the departing vehicle, and the cost-to-go is then equal to the total discounted time spent by all the passengers in the aircraft during its operation.

We now note that $\pi^*(1,y) = 1 \iff J^*(1,T) \leq J^*(1,y)$. First, if $\pi^*(1,y) = 1$, then $J^*(1,y) = \frac{1}{\beta+\mu+\lambda}\left[y + \lambda\frac{y+\gamma}{\beta+\mu} + \mu\frac{\lambda}{\beta+\lambda}\gamma c\right] \geq J^*(1,T) = \frac{\lambda}{\beta+\lambda}\gamma c$. We obtain that $\left(1+\frac{\lambda}{\beta+\mu}\right)y \geq \lambda\gamma c - \frac{\lambda\gamma}{\beta+\mu}$, or $y \geq \frac{\lambda\gamma(c(\beta+\mu)-1)}{\beta+\mu+\lambda}$. Conversely, if $\pi^*(1,y) = 0$, then $J^*(1,y) = \frac{1}{\beta+\lambda}\left[y + \lambda\frac{y+\gamma}{\beta+\mu}\right] < J^*(1,T) = \frac{\lambda}{\beta+\lambda}\gamma c$. We then obtain directly that $\left(1+\frac{\lambda}{\beta+\mu}\right)y < \lambda\gamma c - \frac{\lambda\gamma}{\beta+\mu}$. This proves that $\pi^*(1,y) = 1 \iff y \geq \frac{\lambda\gamma(c(\beta+\mu)-1)}{\beta+\mu+\lambda}$. $\square$

A.3. Proof of Proposition 2

For notational ease, we replace λ_1, γ_{11} and γ_{12} by λ, γ_1 and γ_2, respectively, By definition:

$$J^*(0,T,T) = 0$$
$$J^*(0,T,y_2+\gamma_2) = \frac{1}{\beta+\mu}\left[(y_2+\gamma_2) + \mu\min\{J^*(0,T,y_2+\gamma_2), J^*(0,T,T)\}\right] = \frac{y_2+\gamma_2}{\beta+\mu}$$

$$J^*(0, y_1 + \gamma_1, T) = \frac{1}{\beta + \mu} \left[(y_1 + \gamma_1) + \mu \min \left\{ J^*(0, y_1 + \gamma_1, T), J^*(0, T, T) \right\} \right] = \frac{y_1 + \gamma_1}{\beta + \mu}$$

$$J^*(1, T, T) = \frac{1}{\beta + \lambda} \left[\lambda J^*(0, T, T) + \lambda(\gamma_1 c_1 + \gamma_2 c_2) \right]$$

Let us now derive the optimal in state $(0, y_1 + \gamma_1, y_2 + \gamma_2)$. We have:

$$J^*(0, y_1 + \gamma_1, y_2 + \gamma_2) = \frac{1}{\beta + \mu} \left[(y_1 + \gamma_1) + (y_2 + \gamma_2) + \mu \min \left\{ J^*(0, T, y_2 + \gamma_2), J^*(0, y_1 + \gamma_1, T), J^*(0, y_1 + \gamma_1, y_2 + \gamma_2) \right\} \right]$$

We obtain $\tilde{J}_0(0, y_1 + \gamma_1, y_2 + \gamma_2) = \frac{(y_1 + \gamma_1) + (y_2 + \gamma_2)}{\beta}$, so $\tilde{J}_0(0, y_1 + \gamma_1, y_2 + \gamma_2) \geq J^*(0, y_1 + \gamma_1, T)$ and $\tilde{J}_0(0, y_1 + \gamma_1, y_2 + \gamma_2) \geq J^*(0, T, y_2 + \gamma_2)$. Moreover, $J^*(0, T, y_2 + \gamma_2) \leq J^*(0, y_1 + \gamma_1, T) \iff y_2 + \gamma_2 \leq y_1 + \gamma_1$. Therefore:

$$\pi^*(0, y_1 + \gamma_1, y_2 + \gamma_2) = \begin{cases} 1 & \text{if } y_2 + \gamma_2 \leq y_1 + \gamma_1 \\ 2 & \text{if } y_1 + \gamma_1 \leq y_2 + \gamma_2 \end{cases}$$

We now turn to the optimal policy in states $(1, y_1, T)$ and $(1, T, y_1)$. As in Proposition 1, we obtain:

$$\pi^*(1, y_1, T) = \begin{cases} 1 & \text{if } y_1 \geq \frac{\lambda \gamma_1 (c_1(\beta + \mu) - 1)}{\beta + \mu + \lambda} \\ 0 & \text{otherwise} \end{cases}$$

$$\pi^*(1, T, y_2) = \begin{cases} 2 & \text{if } y_2 \geq \frac{\lambda \gamma_2 (c_2(\beta + \mu) - 1)}{\beta + \mu + \lambda} \\ 0 & \text{otherwise} \end{cases}$$

Finally, we turn to the optimal policy in state $(1, y_1, y_2)$. We have:

$$J^*(1, y_1, y_2) = \frac{1}{\beta + \mu + \lambda} \left[y_1 + y_2 + \lambda J^*(0, y_1 + \gamma_1, y_2 + \gamma_2) + \mu \min \left\{ J^*(1, T, y_2), J^*(1, y_1, T), J^*(1, y_1, y_2) \right\} \right]$$

Without loss of generality, we assume that $y_2 + \gamma_2 \leq y_1 + \gamma_1$. Therefore:

$$J^*(0, y_1 + \gamma_1, y_2 + \gamma_2) = \frac{1}{\beta + \mu} \left[(y_1 + \gamma_1) + (y_2 + \gamma_2) + \frac{\mu}{\beta + \mu} (y_2 + \gamma_2) \right]$$

We distinguish three cases.

Case 1 $y_1 < \frac{\lambda \gamma_1 (c_1(\beta + \mu) - 1)}{\beta + \mu + \lambda}$ *and* $y_2 < \frac{\lambda \gamma_2 (c_2(\beta + \mu) - 1)}{\beta + \mu + \lambda}$

Then, $\pi^*(1, y_1, T) = \pi^*(1, T, y_2) = 0$, and we have:

$$J^*(1, y_1, T) = \frac{1}{\beta + \lambda} \left[y_1 + \lambda \gamma_2 c_2 + \lambda \frac{y_1 + \gamma_1}{\beta + \mu} \right] \qquad \text{and} \qquad J^*(1, T, y_2) = \frac{1}{\beta + \lambda} \left[y_2 + \lambda \gamma_1 c_1 + \lambda \frac{y_2 + \gamma_2}{\beta + \mu} \right]$$

Then, we obtain:

$$J^*(1, T, y_2) \leq J^*(1, y_1, T) \iff \left(1 + \frac{\lambda}{\beta + \mu} \right) y_1 - \lambda \gamma_1 \left(c_1 - \frac{1}{\beta + \mu} \right) \geq \left(1 + \frac{\lambda}{\beta + \mu} \right) y_2 - \lambda \gamma_2 \left(c_2 - \frac{1}{\beta + \mu} \right)$$

$$\iff y_1 - \frac{\lambda \gamma_1 (c_1(\beta + \mu) - 1)}{\beta + \mu + \lambda} \geq y_2 - \frac{\lambda \gamma_2 (c_2(\beta + \mu) - 1)}{\beta + \mu + \lambda}$$

Moreover, $\tilde{J}_0(1, y_1, y_2) = \frac{1}{\beta + \lambda} \left[y_1 + y_2 + \frac{\lambda}{\beta + \mu} \left((y_1 + \gamma_1) + (y_2 + \gamma_2) + \frac{\mu}{\beta + \mu} (y_2 + \gamma_2) \right) \right]$. We therefore obtain:

$$J^*(1, T, y_2) \leq \tilde{J}_0(1, y_1, y_2) \iff \lambda \gamma_1 c_1 \leq y_1 + \frac{\lambda}{\beta + \mu} \left((y_1 + \gamma_1) + \frac{\mu}{\beta + \mu} (y_2 + \gamma_2) \right)$$

$$\iff \left(1 + \frac{\lambda}{\beta + \mu} \right) y_1 - \lambda \gamma_1 \left(c_1 - \frac{1}{\beta + \mu} \right) + \frac{\lambda \mu}{(\beta + \mu)^2} (y_2 + \gamma_2) \geq 0$$

$$\iff y_1 - \frac{\lambda \gamma_1 (c_1(\beta + \mu) - 1)}{\beta + \mu + \lambda} + \frac{\lambda \mu}{(\beta + \mu)(\beta + \mu + \lambda)} (y_2 + \gamma_2) \geq 0$$

Similarly, $J^*(1, y_1, T) \leq \tilde{J}_0(1, y_1, y_2) \iff y_2 - \frac{\lambda \gamma_2 (c_2(\beta + \mu) - 1)}{\beta + \mu + \lambda} + \frac{\lambda \mu}{(\beta + \mu)(\beta + \mu + \lambda)} (y_2 + \gamma_2) \geq 0$.

Thus, $\pi^*(1, y_1, y_2) = 0$ if $y_1 - \frac{\lambda \gamma_1 (c_1(\beta + \mu) - 1)}{\beta + \mu + \lambda} + \frac{\lambda \mu (y_2 + \gamma_2)}{(\beta + \mu)(\beta + \mu + \lambda)} < 0$ and $y_2 - \frac{\lambda \gamma_2 (c_2(\beta + \mu) - 1)}{\beta + \mu + \lambda} + \frac{\lambda \mu (y_2 + \gamma_2)}{(\beta + \mu)(\beta + \mu + \lambda)} < 0$. Otherwise, $\pi^*(1, y_1, y_2) = 1$ if $y_1 - \frac{\lambda \gamma_1 (c_1(\beta + \mu) - 1)}{\beta + \mu + \lambda} \geq y_2 - \frac{\lambda \gamma_2 (c_2(\beta + \mu) - 1)}{\beta + \mu + \lambda}$, and $\pi^*(1, y_1, y_2) = 2$ otherwise.

Case 2 $y_1 \geq \frac{\lambda\gamma_1(c_1(\beta+\mu)-1)}{\beta+\mu+\lambda}$ *and* $y_2 < \frac{\lambda\gamma_2(c_2(\beta+\mu)-1)}{\beta+\mu+\lambda}$

Then, $\pi^*(1,y_1,T)=1$ and $\pi^*(1,T,y_1)=0$ and $J^*(1,T,y_2) \leq J(1,T,T) \leq J(1,y_1,T)$. Moreover, $J^*(1,T,y_2) = \frac{1}{\beta+\lambda}\left[y_2 + \lambda\gamma_1 c_1 + \lambda\frac{y_2+\gamma_2}{\beta+\mu}\right]$ and $\tilde{J}_0(1,y_1,y_2) = \frac{1}{\beta+\lambda}\left[y_1+y_2 + \frac{\lambda}{\beta+\mu}\left((y_1+\gamma_1)+(y_2+\gamma_2)+\frac{\mu}{\beta+\mu}(y_2+\gamma_2)\right)\right]$. As in Case 1:

$$J^*(1,T,y_2) \leq \tilde{J}_0(1,y_1,y_2) \iff y_1 - \frac{\lambda\gamma_1(c_1(\beta+\mu)-1)}{\beta+\mu+\lambda} + \frac{\lambda\mu}{(\beta+\mu)(\beta+\mu+\lambda)}(y_2+\gamma_2) \geq 0$$

This last inequality is satisfied because $y_1 \geq \frac{\lambda\gamma_1(c_1(\beta+\mu)-1)}{\beta+\mu+\lambda}$. Therefore, $\pi^*(1,y_1,y_2)=1$.

Case 3 $y_1 \geq \frac{\lambda\gamma_1(c_1(\beta+\mu)-1)}{\beta+\mu+\lambda}$ *and* $y_2 \geq \frac{\lambda\gamma_2(c_2(\beta+\mu)-1)}{\beta+\mu+\lambda}$

Then, $\pi^*(1,y_1,T)=1$ and $\pi^*(1,T,y_2)=2$, and we have:

$$J^*(1,y_1,T) = \frac{1}{\beta+\mu+\lambda}\left(y_1 + \lambda\frac{y_1+\gamma_1}{\beta+\mu} + \lambda\gamma_2 c_2 + \mu\frac{\lambda}{\beta+\lambda}(\gamma_1 c_1 + \gamma_2 c_2)\right)$$
$$J^*(1,T,y_2) = \frac{1}{\beta+\mu+\lambda}\left(y_2 + \lambda\frac{y_2+\gamma_2}{\beta+\mu} + \lambda\gamma_1 c_1 + \mu\frac{\lambda}{\beta+\lambda}(\gamma_1 c_1 + \gamma_2 c_2)\right)$$

Therefore:

$$J^*(1,T,y_2) \leq J^*(1,y_1,T) \iff y_2 + \lambda\frac{y_2+\gamma_2}{\beta+\mu} + \lambda\gamma_1 c_1 + \mu\frac{\lambda}{\beta+\lambda}(\gamma_1 c_1 + \gamma_2 c_2) \leq y_1 + \lambda\frac{y_1+\gamma_1}{\beta+\mu} + \lambda\gamma_2 c_2 + \mu\frac{\lambda}{\beta+\lambda}(\gamma_1 c_1 + \gamma_2 c_2)$$

$$\iff y_1 - \frac{\lambda\gamma_1(c_1(\beta+\mu)-1)}{\beta+\mu+\lambda} \geq y_2 - \frac{\lambda\gamma_2(c_2(\beta+\mu)-1)}{\beta+\mu+\lambda} J^*(1,T,y_2) \leq \tilde{J}_0(1,y_1,y_2)$$

$$\frac{1}{\beta+\lambda}\left[(y_1+y_2) + \frac{\lambda}{\beta+\mu}\left(y_1+\gamma_1+y_2+\gamma_2 + \frac{\mu}{\beta+\mu}(y_2+\gamma_2)\right)\right]$$

$$\iff \lambda(\beta+\lambda)\gamma_1 c_1 + \lambda\mu(\gamma_1 c_1 + \gamma_2 c_2) \leq (\beta+\mu+\lambda)y_1 + \mu y_2$$

$$+ (\beta+\mu+\lambda)\frac{\lambda}{\beta+\mu}(y_1+\gamma_1) + \frac{\mu\lambda}{\beta+\mu}(y_2+\gamma_2) + \frac{\lambda\mu(\beta+\mu+\lambda)}{(\beta+\mu)^2}(y_2+\gamma_2)$$

$$\iff (\beta+\mu+\lambda)\left[\left(1+\frac{\lambda}{\beta+\mu}\right)y_1 - \lambda\gamma_1\left(c_1 - \frac{1}{\beta+\mu}\right)\right]$$

$$+ \mu\left[\left(1+\frac{\lambda}{\beta+\mu}\right)y_2 - \lambda\gamma_2\left(c_2 - \frac{1}{\beta+\mu}\right)\right] + \frac{\lambda\mu(\beta+\mu+\lambda)}{(\beta+\mu)^2}(y_2+\gamma_2) \geq 0$$

$$\iff (\beta+\mu+\lambda)\left[y_1 - \frac{\lambda\gamma_1(c_1(\beta+\mu)-1)}{\beta+\mu+\lambda}\right] + \mu\left[y_2 - \frac{\lambda\gamma_2(c_2(\beta+\mu)-1)}{\beta+\mu+\lambda}\right] + \frac{\lambda\mu}{\beta+\mu}(y_2+\gamma_2) \geq 0$$

Using the same procedure, we obtain that:

$$J^*(1,T,y_2) \leq \tilde{J}_0(1,y_1,y_2) \iff (\beta+\mu+\lambda)\left[y_2 - \frac{\lambda\gamma_2(c_2(\beta+\mu)-1)}{\beta+\mu+\lambda}\right] + \mu\left[y_1 - \frac{\lambda\gamma_1(c_1(\beta+\mu)-1)}{\beta+\mu+\lambda}\right] + \frac{\lambda\mu}{\beta+\mu}(y_2+\gamma_2) \geq 0$$

The last two inequalities are satisfied since $y_1 \geq \frac{\lambda\gamma_1(c_1(\beta+\mu)-1)}{\beta+\mu+\lambda}$ and $y_2 \geq \frac{\lambda\gamma_2(c_2(\beta+\mu)-1)}{\beta+\mu+\lambda}$. Thus, $\pi^*(1,y_1,y_2)=1$.

In conclusion, we have shown that:

$$\pi^*(1,y_1,y_2) = \begin{cases} 1 & \text{if } y_1 - \frac{\lambda\gamma_1(c_1(\beta+\mu)-1)}{\beta+\mu+\lambda} \geq y_2 - \frac{\lambda\gamma_2(c_2(\beta+\mu)-1)}{\beta+\mu+\lambda} \text{ and } y_1 - \frac{\lambda\gamma_1(c_1(\beta+\mu)-1)}{\beta+\mu+\lambda} + \frac{\lambda\mu(y_2+\gamma_2)}{(\beta+\mu)(\beta+\mu+\lambda)} \geq 0 \\ 2 & \text{if } y_2 - \frac{\lambda\gamma_2(c_2(\beta+\mu)-1)}{\beta+\mu+\lambda} \geq y_1 - \frac{\lambda\gamma_1(c_1(\beta+\mu)-1)}{\beta+\mu+\lambda} \text{ and } y_2 - \frac{\lambda\gamma_2(c_2(\beta+\mu)-1)}{\beta+\mu+\lambda} + \frac{\lambda\mu(y_2+\gamma_2)}{(\beta+\mu)(\beta+\mu+\lambda)} \geq 0 \\ 0 & \text{if } y_1 - \frac{\lambda\gamma_1(c_1(\beta+\mu)-1)}{\beta+\mu+\lambda} + \frac{\lambda\mu(y_2+\gamma_2)}{(\beta+\mu)(\beta+\mu+\lambda)} < 0 \text{ and } y_2 - \frac{\lambda\gamma_2(c_2(\beta+\mu)-1)}{\beta+\mu+\lambda} + \frac{\lambda\mu(y_2+\gamma_2)}{(\beta+\mu)(\beta+\mu+\lambda)} < 0 \end{cases}$$

We proceed similarly in the case where $y_2 + \gamma_2 \geq y_1 + \gamma_1$. This completes the proof. $\square$

A.4. Proof of Proposition 3

We make use of the following lemmas. Lemma 2 expresses the terminal costs incurred once all the departing vehicles have left and some vehicles are yet to arrive. In this case, the cost-to-go is simply given by the costs of misconnections from all arriving vehicles $i \in \mathcal{A}(x)$, discounted, for each vehicle i, by a factor $\frac{\lambda_i}{\beta + \lambda_i}$. Lemma 3 leverages this result to establish the result from Proposition 3 when $|\mathcal{B}(y)| = 1$: when only one vehicle $l \in \mathcal{B}(y)$ is yet to depart, then the optimal policy is to operate vehicle l if and only if $y_l \geq \sum_{i \in \mathcal{A}(x)} \lambda_i \gamma_{il} c_l$.

LEMMA 2. *We denote by $\bar{T}$ the terminal state where $y_j = T$ for all $j = 1, \cdots, m$. For any arrival state x, the cost-to-go in the terminal departure state is given by:*

$$J^*(x, \bar{T}) = \sum_{i \in \mathcal{A}(x)} \frac{\lambda_i}{\beta + \lambda_i} \left(\sum_{j=1}^{p} \gamma_{ij} c_j \right)$$

LEMMA 3. *Let y be such that $y_l \geq 0$ and $y_j = 0, \forall j \neq l$. Then, we have:*

$$\Pi^*(x, y) = \{l\} \quad \text{if } y_l \geq \sum_{i \in \mathcal{A}(x)} \lambda_i \gamma_{il} c_l$$

$$\Pi^*(x, y) = \{0\} \quad \text{otherwise}$$

An overview of the proof of Proposition 3 is shown in Figure 10 with $|\mathcal{A}(x)| = 4$ and $|\mathcal{B}(y)| = 5$. First, Lemma 3 shows the proposition when $|\mathcal{B}(y)| = 1$. We then proceed by double induction over $|\mathcal{A}(x)| \geq 0$ and $|\mathcal{B}(y)| \geq 1$ (indicated by the red numbers in the figure). We first show the proposition when $|\mathcal{A}(x)| = 0$, and then proceed by induction over $|\mathcal{A}(x)|$. For each value of $|\mathcal{A}(x)|$, we proceed by induction over $|\mathcal{B}(y)|$. This latter induction uses Lemma 3 as a starting point, and then proceeds by increasing order of $|\mathcal{B}(y)|$.
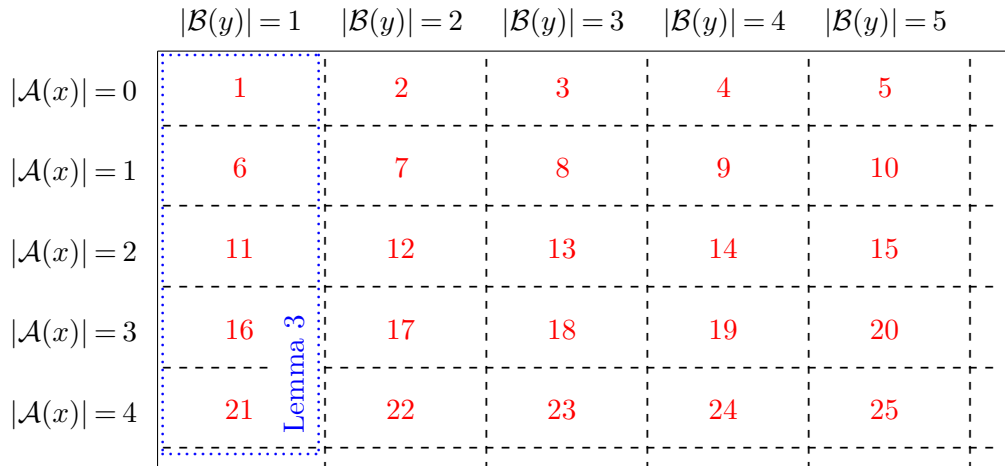
	$\lvert\mathcal{B}(y)\rvert = 1$	$\lvert\mathcal{B}(y)\rvert = 2$	$\lvert\mathcal{B}(y)\rvert = 3$	$\lvert\mathcal{B}(y)\rvert = 4$	$\lvert\mathcal{B}(y)\rvert = 5$
$\lvert\mathcal{A}(x)\rvert = 0$	1	2	3	4	5
$\lvert\mathcal{A}(x)\rvert = 1$	6	7	8	9	10
$\lvert\mathcal{A}(x)\rvert = 2$	11	12	13	14	15
$\lvert\mathcal{A}(x)\rvert = 3$	16	17	18	19	20
$\lvert\mathcal{A}(x)\rvert = 4$	21	22	23	24	25

(Lemma 3)

Figure 10   Overview of the double induction structure in the proof of Proposition 3.

A.4.1. Proof of Lemma 2   We proceed by induction over $|\mathcal{A}(x)|$. The property is trivial when $|\mathcal{A}(x)| = 0$ (the cost-to-go is zero when there is no vehicle). Let us now consider an arrival vector x. Per the induction hypothesis, we know that for any $k \in \mathcal{A}(x)$, we have $J^*(f_k(x, y)) = \sum_{i \in \mathcal{A}(x) \setminus \{k\}} \frac{\lambda_i}{\beta + \lambda_i} \left(\sum_{j=1}^{p} \gamma_{ij} c_j \right)$. Therefore:

$$J^*(x, \bar{T}) = \frac{1}{\beta + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left[\sum_{k \in \mathcal{A}(x)} \lambda_k J^*(f_k(x, y)) + \sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j=1}^{p} \gamma_{ij} c_j \right) \right]$$

$$= \frac{1}{\beta + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left[\sum_{k \in \mathcal{A}(x)} \sum_{i \in \mathcal{A}(x) \setminus \{k\}} \lambda_k \frac{\lambda_i}{\beta + \lambda_i} \left(\sum_{j=1}^{p} \gamma_{ij} c_j \right) + \sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j=1}^{p} \gamma_{ij} c_j \right) \right]$$

$$= \frac{1}{\beta + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left[\sum_{i \in \mathcal{A}(x)} \lambda_i \left(1 + \frac{\sum_{k \in \mathcal{A}(x) \setminus \{i\}} \lambda_k}{\beta + \lambda_i} \right) \left(\sum_{j=1}^{p} \gamma_{ij} c_j \right) \right]$$

$$= \sum_{i \in \mathcal{A}(x)} \frac{\lambda_i}{\beta + \lambda_i} \left(\sum_{j=1}^{p} \gamma_{ij} c_j \right)$$

This completes the proof. $\square$

A.4.2. Proof of Lemma 3 We proceed by induction over $|\mathcal{A}(x)|$. First, let us consider the case where $|\mathcal{A}(x)| = 0$, i.e., all incoming vehicles have already arrived at the facility. Then the proposition simply states that the optimal policy is to operate vehicle l, since there is no benefit of waiting. Formally, we have $J^* (g_l(x,y)) = 0$ and $\tilde{J}_0(x,y) = \frac{y_l}{\beta} \geq 0$, so we obtain directly that $l \in \Pi^*(x,y)$, and $y_l \geq \sum_{i \in \mathcal{A}(x)} \lambda_i \gamma_{il} c_l = 0$.

We now consider an arrival vector x, and assume that the property holds for any arrival vector $\underline{x}$ such that $|\mathcal{A}(\underline{x})| = |\mathcal{A}(x)| - 1$. By definition, we have:

$$\tilde{J}_0(x,y) = \frac{1}{\beta + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left[y_l + \sum_{i \in \mathcal{A}(x)} \lambda_i J^* (f_i(x,y)) + \sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j \neq l} \gamma_{ij} c_j \right) \right]$$

$$J^* (g_l(x,y)) = \sum_{i \in \mathcal{A}(x)} \frac{\lambda_i}{\beta + \lambda_i} \left(\sum_{j=1}^{p} \gamma_{ij} c_j \right) \qquad \text{per Lemma 2}$$

$$J^*(x,y) = \min \left\{ J^* (g_l(x,y)), \tilde{J}_0(x,y) \right\}$$

We distinguish two cases:

Case 1 *There exists $k \in \mathcal{A}(x)$ such that $\Pi^* (f_k(x,y)) = \{0\}$—i.e., $l \notin \Pi^*(f_k(x,y)), \forall l \in \mathcal{B}(y)$*

From the induction hypothesis, we have $y_l \leq \sum_{i \in \mathcal{A}(x) \setminus \{k\}} \lambda_i \gamma_{il} c_l$, which implies that $y_l \leq \sum_{i \in \mathcal{A}(x)} \lambda_i \gamma_{il} c_l$. We aim to show that $\Pi^*(x,y) = \{0\}$, i.e., that $\tilde{J}_0(x,y) \leq J^* (g_l(x,y))$. In other words, if it is optimal to wait after vehicle k has arrived, it is also optimal to wait before vehicle k arrives. For any $m \in \mathcal{A}(x)$, we have:

$$\tilde{J}_0 (f_m(x,y)) = \min \left\{ J^* (g_l (f_m(x,y))), \tilde{J}_0 (f_i(x,y)) \right\}$$

$$\leq J^* (g_l (f_m(x,y))) = J^* (f_m(x,T))$$

$$= \sum_{i \in \mathcal{A}(x) \setminus \{m\}} \frac{\lambda_i}{\beta + \lambda_i} \left(\sum_{j=1}^{p} \gamma_{ij} c_j \right) \qquad \text{per Lemma 2}$$

Therefore:

$$\tilde{J}_0(x,y) = \frac{1}{\beta + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left[y_l + \sum_{m \in \mathcal{A}(x)} \lambda_m J^* (f_m(x,y)) + \sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j \neq l} \gamma_{ij} c_j \right) \right]$$

$$\leq \frac{1}{\beta + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left[y_l + \sum_{m \in \mathcal{A}(x)} \sum_{i \in \mathcal{A}(x) \setminus \{m\}} \frac{\lambda_m \lambda_i}{\beta + \lambda_i} \left(\sum_{j=1}^{p} \gamma_{ij} c_j \right) + \sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j \neq l} \gamma_{ij} c_j \right) \right]$$

$$\leq \frac{1}{\beta + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left[\sum_{i \in \mathcal{A}(x)} \sum_{m \in \mathcal{A}(x) \setminus \{i\}} \frac{\lambda_m \lambda_i}{\beta + \lambda_i} \left(\sum_{j=1}^{p} \gamma_{ij} c_j \right) + \sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j=1}^{p} \gamma_{ij} c_j \right) \right] \qquad \text{because } y_l \leq \sum_{i \in \mathcal{A}(x)} \lambda_i \gamma_{il} c_l$$

$$= \frac{1}{\beta + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left[\sum_{i \in \mathcal{A}(x)} \lambda_i \left(1 + \frac{\sum_{m \in \mathcal{A}(x) \setminus \{k\}} \lambda_m}{\beta + \lambda_i} \right) \left(\sum_{j=1}^{p} \gamma_{ij} c_j \right) \right]$$

$$= \sum_{i \in \mathcal{A}(x)} \frac{\lambda_i}{\beta + \lambda_i} \left(\sum_{j=1}^{p} \gamma_{ij} c_j \right) = J^* \left(g_l(x, y) \right)$$

Case 2 $l \in \Pi^* \left(f_i(x, y) \right)$ *for all* $i \in \mathcal{A}(x)$

We have, for all $i \in \mathcal{A}(x)$, $J^* \left(f_i(x, y) \right) = J^* \left(g_j \left(f_i(x, y) \right) \right) = J^* \left(f_i(x, T) \right)$, so from Lemma 2, it comes: $J^* \left(f_i(x, y) \right) = \sum_{k \in \mathcal{A}(x) \setminus \{i\}} \frac{\lambda_k}{\beta + \lambda_k} \left(\sum_{j=1}^{p} \gamma_{kj} c_j \right)$. Therefore:

$$l \in \Pi^* \left(f_i(x, y) \right) \iff \sum_{i \in \mathcal{A}(x)} \frac{\lambda_i}{\beta + \lambda_i} \left(\sum_{j=1}^{p} \gamma_{ij} c_j \right) \le \frac{1}{\beta + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left[y_l + \sum_{i \in \mathcal{A}(x)} \lambda_i J^* \left(f_i(x, y) \right) + \sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j \neq l} \gamma_{ij} c_j \right) \right]$$

$$\iff \sum_{i \in \mathcal{A}(x)} \frac{\lambda_i}{\beta + \lambda_i} \left(\sum_{j=1}^{p} \gamma_{ij} c_j \right) \le \frac{1}{\beta + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left[y_l + \sum_{i \in \mathcal{A}(x)} \sum_{k \in \mathcal{A}(x) \setminus \{i\}} \frac{\lambda_i \lambda_k}{\beta + \lambda_k} \left(\sum_{j=1}^{p} \gamma_{kj} c_j \right) + \sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j \neq l} \gamma_{ij} c_j \right) \right]$$

$$\iff \sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j=1}^{p} \gamma_{ij} c_j \right) + \sum_{i \in \mathcal{A}(x)} \sum_{k \in \mathcal{A}(x) \setminus \{i\}} \frac{\lambda_k \lambda_i}{\beta + \lambda_i} \left(\sum_{j=1}^{p} \gamma_{ij} c_j \right)$$

$$\le y_l + \sum_{i \in \mathcal{A}(x)} \sum_{k \in \mathcal{A}(x) \setminus \{i\}} \frac{\lambda_i \lambda_k}{\beta + \lambda_k} \left(\sum_{j=1}^{p} \gamma_{kj} c_j \right) + \sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j \neq l} \gamma_{ij} c_j \right)$$

$$\iff \sum_{i \in \mathcal{A}(x)} \lambda_i \gamma_{il} c_l \le y_l$$

This completes the proof. $\square$

A.4.3. Proof of Proposition 3 We proceed by induction over $|\mathcal{A}(x)| \ge 0$ and $|\mathcal{B}(y)| \ge 1$.

First, note that, for $|\mathcal{A}(x)| = 0$, $\mathcal{D}(x, y) = \emptyset$, so the property is equivalent to $\Pi^*(x, y) = \mathcal{B}(y)$ and $J^*(x, y) = 0$. We know through a simple induction over $|\mathcal{B}(y)| \ge 1$ that it is clearly satisfied. Indeed, if $|\mathcal{B}(y)| = 1$, then we can write $y_l \ge 0$ and $y_j = 0, \forall j \neq l$. Then $J^* \left(g_l(x, y) \right)$ and $\tilde{J}_0(x, y) = \frac{y_l}{\beta}$, which clearly implies that $l \in \pi^*(x, y)$ and $J^*(x, y) = 0$. We now consider y such that $|\mathcal{B}(y)| \ge 2$, and assume that the property holds for any $\underline{y}$ such that $\left| \mathcal{B}(\underline{y}) \right| = |\mathcal{B}(y)| - 1$. Then, for each $l \in \mathcal{B}(y)$, we have $J^* \left(g_l(x, y) \right) = 0$ per the induction hypothesis, so $J^* \left(g_l(x, y) \right) \le \tilde{J}_0(x, y) = \frac{\sum_{j \in \mathcal{B}(y)} y_j}{\beta}$, so $l \in \Pi^*(x, y)$ and $J^*(x, y) = 0$. This completes the proof in the case where $|\mathcal{A}(x)| = 0$.

We now consider a departure vector x such that $|\mathcal{A}(x)| \ge 1$ and assume that the property holds in any state $(\underline{x}, y)$ such that $|\mathcal{A}(\underline{x})| = |\mathcal{A}(x)| - 1$. We refer to this induction hypothesis as $(\mathcal{H}_x)$. To prove the property in state (x, y), we proceed by induction over $|\mathcal{B}(y)| \ge 1$.

We first consider a departure vector y such that $|\mathcal{B}(y)| = 1$, i.e., $y_l \ge 0$ and $y_j = 0, \forall j \neq l$. We know from Lemma 3 that $l \in \Pi^*(x, y) \iff y_l \ge \sum_{i \in \mathcal{A}(x)} \lambda_i \gamma_{il} c_l$. Moreover, we have, by denoting by $\bar{T}$ the terminal state where $y_j = T$ for all $j = 1, \cdots, m$:

$$\text{If } y_l \ge \sum_{i \in \mathcal{A}(x)} \lambda_i \gamma_{il} c_l: \quad J^*(x, y) = J^*(x, \bar{T}) = \frac{1}{\beta + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left[\sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j=1}^{p} \gamma_{ij} c_j \right) + \sum_{i \in \mathcal{A}(x)} \lambda_i J^* \left(f_i(x, \bar{T}) \right) \right]$$

$$\text{Otherwise}: \quad J^*(x, y) = \tilde{J}_0(x, y) = \frac{1}{\beta + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left[y_l + \sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j \neq l} \gamma_{ij} c_j \right) + \sum_{i \in \mathcal{A}(x)} \lambda_i J^* \left(f_i(x, y) \right) \right]$$

This completes the proof in the case where $|\mathcal{B}(y)| = 1$, since $\mathcal{D}(x,y) = \emptyset$ if $y_l \geq \sum_{i \in \mathcal{A}(x)} \lambda_i \gamma_{il} c_l$ and $\mathcal{D}(x,y) = \{l\}$ otherwise.

We now turn to the case where $|\mathcal{B}(y)| \geq 1$, and we assume that the property is satisfied in any state $(x, \underline{y})$ such that $|\mathcal{B}(\underline{y})| = |\mathcal{B}(y)| - 1$. We refer to this induction hypothesis as $(\mathcal{H}_y)$. By definition, we have:

$$\tilde{J}_0(x,y) = \frac{1}{\beta + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left[\sum_{j \in \mathcal{B}(y)} y_j + \sum_{i \in \mathcal{A}(x)} \lambda_i J^*\left(f_i(x,y)\right) + \sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j \notin \mathcal{B}(y)} \gamma_{ij} c_j \right) \right]$$

$$J^*(x,y) = \min\left\{ \min_{l \in \mathcal{B}(y)} J^*\left(g_l(x,y)\right), \tilde{J}_0(x,y) \right\}$$

For any $l \in \mathcal{B}(y)$, we express $J^*\left(g_l(x,y)\right)$ from the induction hypothesis $(\mathcal{H}_y)$ as follows:

- If $l \in \mathcal{D}(x,y)$, we have $\mathcal{D}\left(g_l(x,y)\right) = \mathcal{D}(x,y) \setminus \{l\}$, which yields:

$$J^*\left(g_l(x,y)\right) = \frac{1}{\beta + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left\{ \left(\sum_{j \in \mathcal{D}(x,y)} y_j - y_l \right) + \sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j \notin \mathcal{D}(x,y)} \gamma_{ij} c_j + \gamma_{il} c_l \right) + \sum_{i \in \mathcal{A}(x)} \lambda_i J^*\left(f_i(x, \delta\left(g_l(x,y)\right))\right) \right\}$$

- If $l \notin \mathcal{D}(x,y)$, we have $\mathcal{D}\left(g_l(x,y)\right) = \mathcal{D}(x,y)$ and $\delta\left(g_l(x,y)\right) = \delta(x,y)$, which yields:

$$J^*\left(g_l(x,y)\right) = \frac{1}{\beta + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left\{ \sum_{j \in \mathcal{D}(x,y)} y_j + \sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j \notin \mathcal{D}(x,y)} \gamma_{ij} c_j \right) + \sum_{i \in \mathcal{A}(x)} \lambda_i J^*\left(f_i(x, \delta(x,y))\right) \right\}$$

We now proceed to prove the property from the proposition. We distinguish two cases.

Case 1 $\mathcal{D}(x,y) = \mathcal{B}(y)$

We show that, in this case, $\Pi^*(x,y) = \{0\}$. Let us consider $l \in \mathcal{B}(y)$. Note that, since $\mathcal{D}(x,y) = \mathcal{B}(y)$, we have $\delta(x,y) = (x,y)$ and $(x, \delta\left(g_l(x,y)\right) = g_l(x,y)$. It comes:

$$\tilde{J}_0(x,y) < J^*\left(g_l(x,y)\right)$$

$$\iff \frac{1}{\beta + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left[\sum_{j \in \mathcal{B}(y)} y_j + \sum_{i \in \mathcal{A}(x)} \lambda_i J^*\left(f_i(x,y)\right) + \sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j \notin \mathcal{B}(y)} \gamma_{ij} c_j \right) \right] <$$

$$\frac{1}{\beta + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left\{ \left(\sum_{j \in \mathcal{B}(y)} y_j - y_l \right) + \sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j \notin \mathcal{B}(y)} \gamma_{ij} c_j + \gamma_{il} c_l \right) + \sum_{i \in \mathcal{A}(x)} \lambda_i \underbrace{J^*\left(f_i(x, \delta\left(g_l(x,y)\right))\right)}_{= J^*(g_l(f_i(x,y)))} \right\}$$

$$\iff \underbrace{\left(\sum_{i \in \mathcal{A}(x)} \lambda_i \gamma_{il} c_l - y_l \right)}_{> 0 \text{ because } l \in \mathcal{D}(x,y)} + \sum_{i \in \mathcal{A}(x)} \lambda_i \underbrace{\left[J^*\left(g_l\left(f_i(x,y)\right)\right) - J^*\left(f_i(x,y)\right) \right]}_{\geq 0 \text{ (see below)}} > 0$$

This last inequality is satisfied because $J^*\left(f_i(x,y)\right) = \min\left\{ J^*\left(g_l\left(f_i(x,y)\right)\right), \tilde{J}_0\left(f_i(x,y)\right) \right\} \leq J^*\left(g_l\left(f_i(x,y)\right)\right)$. This shows that $\Pi^*(x,y) = \{0\}$. Moreover, we have $J^*(x,y) = \tilde{J}_0(x,y)$, i.e.:

$$J^*(x,y) = \frac{1}{\beta + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left[\sum_{j \in \mathcal{B}(y)} y_j + \sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j \notin \mathcal{B}(y)} \gamma_{ij} c_j \right) + \sum_{i \in \mathcal{A}(x)} \lambda_i J^*\left(f_i(x,y)\right) \right]$$

This completes the proof in the case where $\mathcal{D}(x,y) = \mathcal{B}(y)$

Case 2 $\mathcal{D}(x,y) \neq \mathcal{B}(y)$

We consider $l \in \mathcal{B}(y) \setminus \mathcal{D}(x,y)$, and show that $l \in \Pi^*(x,y)$. We will show, first, that $J^*(g_l(x,y)) \leq J^*(g_m(x,y))$ for any $m \in \mathcal{B}(y)$, and, second, that $J^*(g_l(x,y)) \leq \tilde{J}_0(x,y)$.

Let us consider $m \in \mathcal{B}(y)$. Note, first, that $J^*(g_l(x,y)) = J^*(g_m(x,y))$ if $m \notin \mathcal{D}(x,y)$, because the expression of $J^*(g_l(x,y))$ does not depend on l when $l \notin \mathcal{D}(x,y)$. We now assume that $m \in \mathcal{D}(x,y)$. We have:

$$J^*(g_l(x,y)) < J^*(g_m(x,y))$$

$$\iff \frac{1}{\beta + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left\{ \sum_{j \in \mathcal{D}(x,y)} y_j + \sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j \notin \mathcal{D}(x,y)} \gamma_{ij} c_j \right) + \sum_{i \in \mathcal{A}(x)} \lambda_i J^*(f_i(x, \delta(x,y))) \right\} <$$

$$\frac{1}{\beta + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left\{ \left(\sum_{j \in \mathcal{D}(x,y)} y_j - y_m \right) + \sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j \notin \mathcal{D}(x,y)} \gamma_{ij} c_j + \gamma_{im} c_m \right) + \sum_{i \in \mathcal{A}(x)} \lambda_i \underbrace{J^*(f_i(x, \delta(g_m(x,y))))}_{=J^*(g_m(f_i(x,y)))} \right\}$$

$$\iff \underbrace{\left(\sum_{i \in \mathcal{A}(x)} \lambda_i \gamma_{im} c_m - y_m \right)}_{>0 \text{ because } m \in \mathcal{D}(x,y)} + \sum_{i \in \mathcal{A}(x)} \lambda_i \underbrace{[J^*(g_m(f_i(x,y))) - J^*(f_i(x,y))]}_{\geq 0 \text{ (as in Case 1)}} > 0$$

We therefore obtain that $J^*(g_l(x,y)) < J^*(g_m(x,y))$ if $m \in \mathcal{D}(x,y)$. This shows that $J^*(g_l(x,y)) \leq J^*(g_m(x,y))$ for all $m \in \mathcal{B}(y)$. Finally, we have:

$$J^*(g_l(x,y)) \leq \tilde{J}_0(x,y)$$

$$\iff \frac{1}{\beta + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left\{ \sum_{j \in \mathcal{D}(x,y)} y_j + \sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j \notin \mathcal{D}(x,y)} \gamma_{ij} c_j \right) + \sum_{i \in \mathcal{A}(x)} \lambda_i J^*(f_i(x, \delta(x,y))) \right\} \leq$$

$$\frac{1}{\beta + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left[\sum_{j \in \mathcal{B}(y)} y_j + \sum_{i \in \mathcal{A}(x)} \lambda_i J^*(f_i(x,y)) + \sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j \notin \mathcal{B}(y)} \gamma_{ij} c_j \right) \right]$$

$$\iff \sum_{j \in \mathcal{B}(y) \setminus \mathcal{D}(x,y)} \underbrace{\left(y_j - \sum_{i \in \mathcal{A}(x)} \lambda_i \gamma_{ij} c_j \right)}_{\geq 0 \text{ for } j \notin \mathcal{D}(x,y)} + \sum_{i \in \mathcal{A}(x)} \lambda_i \underbrace{[J^*(f_i(x,y)) - J^*(f_i(x, \delta(x,y)))]}_{=0 \text{ per the induction hypothesis } (\mathcal{H}_x)} \geq 0$$

This shows that $l \in \Pi^*(x,y)$. As a result, we have $J^*(x,y) = J^*(g_l(x,y))$, which proves that

$$J^*(x,y) = \frac{1}{\beta + \sum_{i \in \mathcal{A}(x)} \lambda_i} \left\{ \left(\sum_{j \in \mathcal{D}(x,y)} y_j - y_l \right) + \sum_{i \in \mathcal{A}(x)} \lambda_i \left(\sum_{j \notin \mathcal{D}(x,y)} \gamma_{ij} c_j + \gamma_{il} c_l \right) + \sum_{i \in \mathcal{A}(x)} \lambda_i J^*(f_i(x, \delta(g_l(x,y)))) \right\}$$

This completes the proof. $\quad\square$

Appendix B: Proof of Statements from Section 3

B.1. Proof of Proposition 4

Let us consider an optimal solution $(\mathcal{S})$ to (GDP-PAX) and assume by contradiction that $\lambda_{rj} = 1$ and $\sum_{t \in \mathcal{T}} w^a_{f_1(r),t} - \sum_{t \in \mathcal{T}} w^d_{jt} \geq \sigma_{rj}$ for some $r \in \mathcal{P}_C$ and $j \in \mathcal{R}_r$. First, note that if $j = f_2(r)$, then Solution $(\mathcal{S})$ violates Constraint (12). We thus assume that $j \neq f_2(r)$ and build a solution $(\widehat{\mathcal{S}})$ such that: (i) $\widehat{w}^a = w^a$ and $\widehat{w}^d = w^d$ (i.e., flight operations remain unchanged), (ii) $\widehat{z} = z$ and $\widehat{z}^S = z^S$ (i.e., passenger accommodations remain unchanged), (iii) $\widehat{\lambda}_{rj} = 0$ and $\widehat{\lambda}_{pi} = \lambda_{pi}$ for each $(p,i) \in \mathcal{P}_C \times \mathcal{F} \setminus \{(r,j)\}$. By construction, Constraint (11) is satisfied because $\sum_{t \in \mathcal{T}} w^a_{f_1(r),t} - \sum_{t \in \mathcal{T}} w^d_{jt} \geq \sigma_{rj}$. Note that Constraint (13) gets relaxed,

and Constraints (12) and (14) do not get modified because $f_2(r) \neq j$. Therefore $(\widehat{S})$ is a feasible solution and, since $\sum_{p \in \mathcal{P}_C} \left(\sum_{i \in \mathcal{R}_p} \delta_{pi}^C \widehat{z}_{pi} + \delta_p^S \widehat{z}_p^S \right) = \sum_{p \in \mathcal{P}_C} \left(\sum_{i \in \mathcal{R}_p} \delta_{pi}^C z_{pi} + \delta_p^S z_p^S \right)$, it is also optimal.

If Constraint (12) is omitted, then we provide in Table 9 a counter-example where the property is not necessarily satisfied. Let us consider the following set of five flights: two incoming flights into Airport C, and three outgoing flights on the segment between Airport C and Airport D. We also consider two passenger itineraries, referred to as p_1 and p_2: p_1 involves a connection from flight 1 to flight 4, and p_2 involves a connection from flight 2 to flight 3 (i.e., $f_1(p_1) = 1$, $f_2(p_1) = 4$, $f_1(p_2) = 2$ and $f_2(p_2) = 3$). We also assume that each of these two itineraries are booked by five passengers (i.e., $n_{p_1} = n_{p_2} = 5$), that each of the three outgoing flights from Airport C have an effective capacity of 5 (i.e., $\Omega_3 = \Omega_4 = \Omega_5 = 5$), and that the minimal connecting time between any pair of flights at airport C is 45 minutes (i.e., $\sigma_{p,i} = 3$ 15-minute periods for each $p \in \{p_1, p_2\}$ and $i \in \{1, 2, 3\}$). Since flight 2 is delayed by 30 minutes (e.g., due to upstream congestion at Airport B), the connection between flight 2 and flight 3 is no longer feasible and itinerary p_2 is therefore disrupted. This implies per Constraint (11) that $\lambda_{p_2,3} = 1$. However, passengers on itinerary p_2 can still make the connection to flight 4, i.e., $\lambda_{p_2,4} = 0$. Since $\lambda_{p_1,3} = \lambda_{p_1,4} = 0$, we have per Constraint (14) $z_{p_1,4} = 5$. The aircraft used to fly flight 4 is therefore at capacity, and passengers on itinerary p_2 cannot be re-accommodated on flight 4. Therefore, they will be re-accommodated on flight 5 and incur a connection delay of 8 hours. However, this is not the optimal solution. Instead, in the absence of Constraint (12), one can set $\lambda_{p_1,4} = 1$ (even though the connection remains feasible), and thus $z_{p_1,4} = 0$ and $z_{p_1,3} = z_{p_2,4} = 5$. Then the passengers on itinerary p_1 incur a delay of 0 and the passengers on itinerary p_2 incur a delay of 1 hour.

Table 9   **Counter-example if Constraint** (12) **is omitted.**

Flight ID	Origin	Destination	Planned Dep.	Planned Arr.	Actual Dep.	Actual Arr.	Eff. Capacity
1	Airport A	Airport C	7:15 am	9:15 am	7:15 am	9:15 am	0
2	Airport B	Airport C	8:15 am	9:15 am	8:45 am	9:45 am	0
3	Airport C	Airport D	10:00 am	11:00 am	10:00 am	11:00 am	5
4	Airport C	Airport D	11:00 am	12:00 am	11:00 am	12:00 am	5
5	Airport C	Airport D	6:00 pm	7:00 pm	6:00 pm	7:00 pm	5

Therefore, the inclusion of Constraint (12) enforces that an itinerary-flight pairing is marked as disrupted *if and only if* the time between the arrival time of the first-leg flight of the itinerary and the departure time of the flight under consideration is insufficient for passengers to connect. This completes the proof. $\square$

B.2. Proof of Proposition 5

First, we note that for any *positive* variables $\tilde{z}$ and $\tilde{\lambda}$ such that $\tilde{z}_{pr} \leq n_p(1 - \tilde{\lambda}_r)$ for each $p, r \in \mathcal{P}_C$, we have $\sum_{p \in \mathcal{P}_C} \tilde{z}_{pr} \leq \left(\sum_{p \in \mathcal{P}_C} n_p \right)(1 - \tilde{\lambda}_r)$. Therefore, the constraint $\tilde{z}_{pr} \leq n_p(1 - \tilde{\lambda}_r), \forall p, r \in \mathcal{P}_C$ defines a tighter feasible region than the constraint $\sum_{p \in \mathcal{P}_C} \tilde{z}_{pr} \leq \left(\sum_{p \in \mathcal{P}_C} n_p \right)(1 - \tilde{\lambda}_r), \forall r \in \mathcal{P}_C$. Second, we note that both constraints define the same feasible region when $\tilde{z}$ is an integer variable and $\tilde{\lambda}$ is a binary variable. To show this, let us consider $r \in \mathcal{P}_C$ and assume that $\sum_{p \in \mathcal{P}_C} \tilde{z}_{pr} \leq \left(\sum_{p \in \mathcal{P}_C} n_p \right)(1 - \tilde{\lambda}_r)$. If $\tilde{\lambda}_r = 1$, this implies that $\sum_{p \in \mathcal{P}_C} \tilde{z}_{pr} = 0$, hence $\tilde{z}_{pr} = 0$ for all $p \in \mathcal{P}_C$. Conversely, if $\tilde{\lambda}_r = 0$, then we already know that $\tilde{z}_{pr} \leq n_p \tilde{\lambda}_p$ for all $p \in \mathcal{P}_C$ because $\tilde{z}_{pr} \leq n_p$ for all $p \in \mathcal{P}_C$ (by assumption). Therefore, if $\sum_{p \in \mathcal{P}_C} \tilde{z}_{pr} \leq \left(\sum_{p \in \mathcal{P}_C} n_p \right)(1 - \tilde{\lambda}_r), \forall r \in \mathcal{P}_C$, then $\tilde{z}_{pr} \leq n_p(1 - \tilde{\lambda}_r), \forall p, r \in \mathcal{P}_C$. This completes the proof. $\square$