OPERATIONS RESEARCH
DOI 10.1287/opre.1080.0630ec
pp. ec31–ec39

e-companion

ONLY AVAILABLE IN ELECTRONIC FORM

Electronic Companion—"Analysis of Airplane Boarding Times" by Eitan Bachmat, Daniel Berend, Luba Sapir, Steven Skiena, and Natan Stolyarov, *Operations Research*, DOI 10.1287/opre.1080.0630.

Online Appendix A: An example of the boarding process model

We provide an example of the boarding process as modelled in Section 2. Consider an airplane with a single passenger per row. Let $W = 2/3$ and $l = 1$. We index passengers according to their position in the queue. The rows of the passengers in the queue are given by the sequence: 5,10,9,11,7,8,6,2,3,4,1, that is, the first passenger to board is in row 5, the next in row 10 and so on. The obvious blocking relations are that passenger i blocks passenger j if $i < j$ and $r_i < r_j$. For example, the first passenger in row 5 blocks the fourth passenger who is trying to get to row 11. What is less obvious, is that the first passenger also blocks passenger 8, who is trying to get to the second row. The reason is that passengers 2 through 7 will be lined up behind the first passenger, since they are all trying to reach rows which are beyond the fifth. Each of these six passengers occupies 2/3 of the distance between successive rows, for a total backlog of 4 rows. Since the difference between the fifth and second row is only three, the eighth passenger has to wait for the first passenger to sit down before he/she can actually reach his/her row. In terms of (1), this says that $5 - 6(2/3) < 2$. Similarly, the eighth passenger blocks the eleventh passenger, who is seated in the first row, via the backlog caused by the presence of the ninth and tenth passengers. These are the only non-trivial blocking relations in this example. The longest chain of blocking relations is of length 4 and is given by passengers 1,8,9,10. The (q, r) coordinate representation of the boarding queue is displayed in Figure 5. The figure also includes a drawing of the maximal chain.

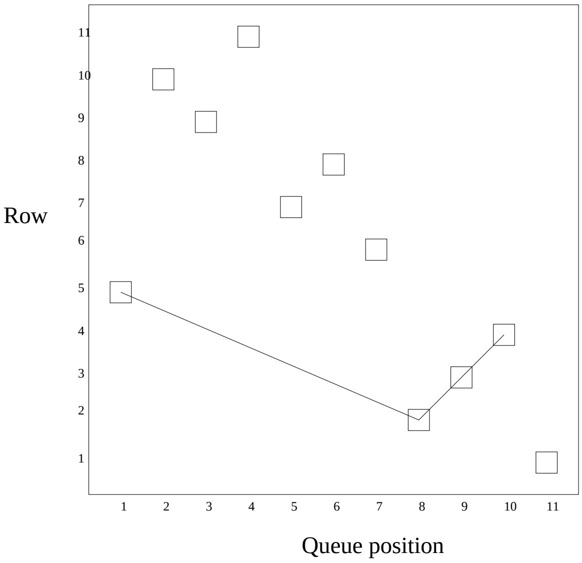

Figure 5: A graphic representation of the queue and maximal chain.

Online Appendix B: Table of permutations

We provide a table of the permutations which were used by Van Landeghem and Beuselinck in their experiments.

Index	Permutation
0	12, 10, 8, 11, 9, 7, 6, 4, 2, 5, 3, 1
1	4, 1, 2, 3
2	12, 4, 8, 5, 9, 1, 11, 3, 7, 6, 10, 2
3	12, 9, 11, 8, 10, 7, 6, 3, 5, 2, 4, 1
4	8, 6, 7, 5, 4, 2, 3, 1
5	8, 3, 6, 1, 4, 7, 2, 5
6	6, 4, 5, 3, 1, 2
7	10, 5, 9, 4, 8, 3, 7, 2, 6, 1
8	2, 3, 1
9	4, 2, 3, 1
10	4, 1, 3, 2
11	6, 3, 5, 2, 4, 1
12	6, 4, 2, 5, 3, 1
13	6, 2, 5, 1, 4, 3
14	10, 8, 6, 4, 2, 9, 7, 5, 3, 1

Table 2: Permutations

Online Appendix C: Boarding time formulas

In addition to the boarding time formulas given in (17), (18), (26), (27), for the F_1 and $F_{2,\sigma}$ policies, we compute boarding times for policies with more than two groups.

Computing $B(F_v, k)$

Consider the case of a uniform back-to-front policy with $v > 2$. As for F_2, it is more convenient to consider the expanded square $[0, v] \times [0, v]$ instead of the unit square. We have the sub-squares $S_i = [i-1, i] \times [m-i, m-i+1]$, $i = 1, \ldots, v$, which lie along the anti-diagonal. The part of the maximal length curve U which contributes to the length is composed as before of segments U_i contained in S_i. In addition, there are legitimate line segments, joining the endpoint of U_i to the initial point of U_{i+1}, which lie below S_i and do not contribute to the length. The arguments presented for the case of F_2 also show that $(0, v-1)$ is the initial point of U (and U_1) and that the initial point of U_i is on the left edge of S_i for all i. In addition, for $i = 1, \ldots, v-1$, the endpoint of U_i lies on the bottom edge of S_i. Thus, the initial point of U_i has the form $(i-1, v-i+\delta_i)$ and the endpoints have the form $(i-1+\beta_i, v-i)$ for some δ_i and β_i. The maximality of U, together with the requirement that U be legitimate, imply the relation

$$\beta_i = 1 - \frac{1-\delta_{i+1}}{k}. \tag{36}$$

Consider U_1 and U_2. The initial point and the endpoint of U_1 lie on the bottom edge of S_1, and hence, by maximality, all of U_1 lies on the bottom edge. Let $(1, v-2+\delta_2)$ be the initial point of U_2 and $(1+\beta_2, v-2)$ its endpoint. Given the endpoint of U_1, we can optimize the location of the initial point of U_2. Depending on δ_2, U_2 may either be a solution to the Euler-Lagrange equation which lies in the interior of S_2 or be composed of a solution which is tangent to the bottom edge, followed by a segment along the bottom edge. In the first case the length of the segment is

$$L(U_2) = \frac{\sqrt{e^{k\beta_2}-1}}{\sqrt{k}} \cdot \frac{\sqrt{(1-\delta_2)e^{k\beta_2}-1}}{\sqrt{e^{k\beta_2}-1}}. \tag{37}$$

We may add the contribution of $L(U_1)$ and differentiate. The sum of contributions is maximized when

$$\delta_2 = -\left(\frac{e^{k\beta_2}}{4} + \frac{4}{e^{k\beta_2}}\right) + \frac{5}{4}\,. \tag{38}$$

Writing $x = \frac{e^{k\beta_1}}{4}$, we see that $\delta_2 = -(x + 1/x) + 5/4$, and, since $x + 1/x \geq 2$ for all $x > 0$, the optimal δ_2 is negative. Hence the U_2 component of the maximal curve is composed of a solution tangent to the bottom of S_2, followed (possibly) by a segment along the bottom. Such curves were already analyzed in the F_2 case. The optimal value for δ_2 is $1/4$ if it is in the range

$$\left(\ln \frac{1}{1-\sqrt{\delta_2}}\right)/k \leq \beta_2. \tag{39}$$

Let $k \geq 3/4 + \ln 2$. Consider the curve U with $\delta_i = 1/4$ for $i = 1, \ldots, v-1$. By (36), this corresponds to $\beta_i = 1 - \frac{3}{4k}$. We note that the analogue of (39) with δ_2 and β_2 replaced by $1/4$ and $1 - \frac{3}{4k}$, respectively, is satisfied when $k \geq 3/4 + \ln 2$. We claim that U is maximal. Assume to the contrary that U' with parameters β'_i is maximal. Let j be the first index for which $\beta'_j \neq 1 - \frac{3}{4k}$. We claim that $\beta'_j < 1 - \frac{3}{4k}$, as otherwise, by the calculations for U_1 and U_2, the path U'_{j+1} from $(j, v-(j+1)+\delta'_{j+1})$

to $(j+\beta'_{j+1}, v-(j+1))$ is not optimal, since $\delta'_{j+1} \geq 1/4$. For the same reason we must have $\beta'_i < 1-\frac{3}{4k}$ for all $i > j$. However, by the computation for F_2 we know that for $k \geq 2\ln 2$ the contribution of the union of U_{v-1} and U_v is not optimal. Since $3/4 > \ln 2$, we are done.

The length of the curve U is

$$B(F_v, k) = \sqrt{vk} - \frac{v-2}{\sqrt{vk}}(\ln 2 + 1/4) - \frac{2\ln 2 - 3/4}{\sqrt{vk}}. \tag{40}$$

The maximal curve for F_3 with $k = 4$ is depicted in Figure 6. The maximal curve for F_v has the same structure with the segments $U_2, \ldots, U_{v-1}$ repeating.

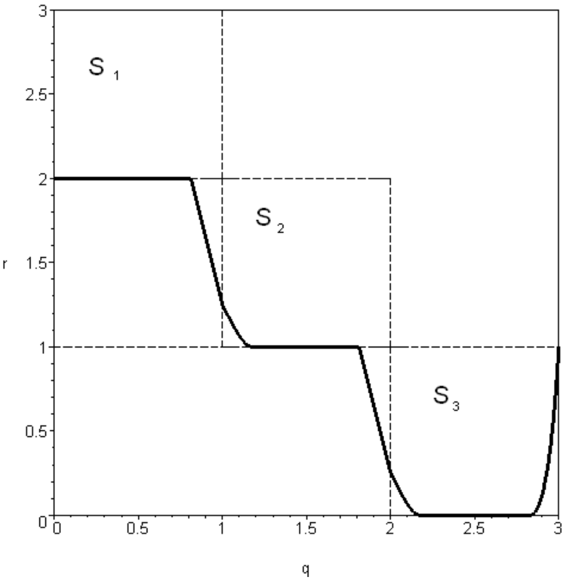

Figure 6: Maximal curve for F_3 with $k = 4$.

Computing $B(F_{v,\sigma}, k)$

We compute $B(F_{v,\sigma}, k)$ under some assumptions on σ and k. Using the representation of passengers in the expanded square $[0, v] \times [0, v]$, passengers in the group $G_{\sigma(i)}$ will be represented by points in the square $S_{i,\sigma}$, consisting of points of the form $i-1 \leq q \leq i$ and $\sigma(i) - 1 \leq r \leq \sigma(i)$. Let U be a maximal curve, composed of curves U_i in $S_{i,\sigma}$. The sequence $\sigma(1), \sigma(2), \ldots, \sigma(v)$ decomposes into blocks Z_j, $1 \leq j \leq \mu(\sigma)$, of size z_j of decreasing sequences, that is, $\sigma(1) > \ldots > \sigma(z_1) < \sigma(z_1+1) > \ldots > \sigma(z_1+z_2) < \sigma(z_1+z_2+1)\ldots$

Let $c_j = \sum_{i=1}^{j} z_i$ be the sequence of partial sums of the sequence z_j. Define the *excess descent* of block Z_j by $e_j = \sigma(c_j) - \sigma(c_{j-1}) - z_j + 1$ if $z_j > 1$ and $e_j = 0$ if $z_j = 1$. Let $e = \sum_j e_j$. Let $\chi(\sigma)$ be the maximum over all decreasing pairs of consecutive elements, $\sigma(i) < \sigma(i+1)$, of $\sigma(i+1) - \sigma(i)$. If

$$k \geq 3/4 + \ln 2 + \lambda(\sigma) - 1, \tag{41}$$

then

$$B(F_{v,\sigma}, k) = \sum_{j=1}^{\mu(\sigma)} B(F_{z_j}, k)\frac{\sqrt{z_j}}{\sqrt{v}} - e_j\frac{1}{\sqrt{kv}}. \tag{42}$$

To see this, consider a maximal curve U in the union of $S_{i,\sigma}$, composed of curves U_i in the respective squares, and legitimate linear segments joining them. As in the previous computations, it is easy to verify that the endpoint of U_i is on the bottom edge of $S_{i,\sigma}$, say at $(i-1+\beta_{i,\sigma}, \sigma(i)-1)$, and the initial point of U_i is on the left border of $S_{i,\sigma}$, say at $(i-1, \sigma(i)-1+\delta_{i,\sigma})$. The endpoint of U_i must be connected to the initial point of U_{i+1} by a legitimate linear segment. If $\sigma(i) < \sigma(i+1)$, then the line segment between any point of $S_{i,\sigma}$ and any point of $S_{i+1,\sigma}$, is legitimate. Therefore we can make independent computations for finding the curves U_i, $\sigma(i) \in Z_j$, for each fixed j, as long as all the curves U_i are non-empty. By (4), when $\sigma(i) > \sigma(i+1)$ we have

$$\sigma(i) - \sigma(i+1) - \delta_{i+1,\sigma} \leq k(1-\beta_{i,\sigma}). \tag{43}$$

The process of identifying a maximal curve for each j separately is the same as that of identifying the maximal curve for the policy F_{z_j}, subject to the replacement of (36) by (43) and multiplication of the density p by the factor z_j/v. As in the case of F_{z_j}, the curve with $\delta_{i,\sigma} = 1/4$, for $c_{j-1}+2 \leq i \leq c_j$, is optimal when it can be constructed subject to (43), which holds when $k \geq \sigma(i)-\sigma(i+1)-1+3/4+\ln 2$ for all i, or equivalently when $k \geq 3/4+\ln 2+\chi(\sigma)-1$. The change in density yields a comparison with $B(F_{z_j}, k)\frac{\sqrt{z_j}}{\sqrt{v}}$. A comparison of (36) and (43) shows that the portion of the curve U_i on the bottom edge of $S_{i,\sigma}$ is shorter (in the standard Euclidean sense) by $\frac{\sigma(i+1)-\sigma(i)-1}{k}$ than the corresponding curve in the computation of $B(F_{z_j}, k)$. Summing over i such that $\sigma(i) \in Z_j$ and taking the density value v into account, we obtain the difference term $-e_j\frac{1}{\sqrt{kv}}$, thus establishing (42).

Computing $B(F^2_{v,\sigma})$ for cyclic permutations

We compute estimated boarding times for half-row policies with a cyclic permutation. We refer the reader to Subsection 4.3 for the various definitions.

Consider the restriction of the boarding process to the first half of the queue. We wish to compare this part of the boarding process to the boarding process associated with the policy $F_{w,\rho}$. The number of passengers is $n/2$ instead of n. Also, h should be replaced by $h/2$, since only half the passengers in each row (those to the right of the aisle) are boarding. This leads to the replacement of k by $k/2$. The order in which row blocks board is by definition the same as that of $F_{w,\rho}$. We conclude that the boarding time for the first half of the queue B_1 is given by

$$B_1 = B(F_{w,\rho}, k/2)/\sqrt{2}\,. \tag{44}$$

By symmetry, the same formula would hold for the boarding time if we considered only the second half of the queue. Assume that $\sigma(1) < \sigma(w)$. This means that each passenger from the last group of the first half of the queue blocks each passenger from the first group in the second half of the queue. Assume furthermore that k is large enough so that the maximal curve U for half the queue contains non-empty segments in the first and last groups, namely U_1 and U_w are non-empty when considering the policy $F_{w,\rho}$ with congestion parameter $k/2$. We conclude that in such cases the maximal curves for the first and second half of the queue can be concatenated to produce a maximal curve of twice the length. Therefore, under such circumstances

$$B(F^2_{w,\sigma}, k) = 2B_1 = \sqrt{2}B(F_{w,\rho}, k/2). \tag{45}$$

In particular, for a back-to-front cyclic permutation,

$$B(F^2_w, k) = \sqrt{2}B(F_w, k/2). \tag{46}$$

Online Appendix D: Boarding time computations

We provide a brief explanation for the various calculations which were used in the computation of the table in Section 6. We first recall the main formulas needed for the computations.

A) The basic formula for F_1 with $k > \ln 2$:

$$B(F_1, k) = \sqrt{k} + \frac{1 - \ln 2}{\sqrt{k}}\,. \tag{47}$$

B) The formula for F_2:

$$B(F_2, k) = \begin{cases} \sqrt{\frac{1}{2k}}(k + \frac{e^k - 1}{4}), & 1 \le k \le 2\ln 2, \\ \sqrt{2k} + \frac{3/4 - 2\ln 2}{\sqrt{2k}}, & k \ge 2\ln 2. \end{cases} \tag{48}$$

C) The formula for F_v, $v > 2$, and $k \ge 3/4 + \ln 2$:

$$B(F_v, k) = \sqrt{vk} - \frac{v-2}{\sqrt{vk}}(\ln 2 + 1/4) - \frac{2\ln 2 - 3/4}{\sqrt{vk}}\,. \tag{49}$$

D) The formula for $F_{v,\sigma}$, where $k \ge 3/4 + \ln 2 + \lambda(\sigma) - 1$:

$$B(F_v, \sigma, k) = \sum_{j=1}^{\mu(\sigma)} B(F_{z_j}, k)\frac{\sqrt{z_j}}{\sqrt{v}} - e_j \frac{1}{\sqrt{kv}}\,. \tag{50}$$

E) The formula for $F^2_{w,\sigma}$:

$$B(F^2_{w,\sigma}, k)\sqrt{n} = \sqrt{2}B(F_{w,\sigma}, k/2)\,. \tag{51}$$

All policies not involving a permutation σ are computed directly from these formulas by setting $k = 4$ and dividing by the expression for F_1, which yields the value 2.15. Among policies which do involve a permutation σ, all the computations are straightforward using a combination of formulas (A-E) above, with the exception of the policies $F^2_{6,\sigma_0}, F^2_{6,\sigma_2}, F_{10,\sigma_7}$ and $F^2_{6,\sigma_{13}}$, for which we proceed as follows:

F_{10,σ_7}: In this case (50) does not hold since some of the U_i's are empty. To compute a maximal curve U, we need to decide which of the U_i's are non-empty. This can be done using a dynamic programming approach, which is easily carried out by hand for small-sized permutations. The maximal curve is obtained when U_i is non-empty for all even values of i. We note that, for even i, the top right corner of $S_{i,\sigma}$ blocks the bottom right corner of $S_{i+2,\sigma}$. As a result, $B(F_{10,\sigma_7}) = \frac{5}{\sqrt{10}}B(F_1, k)$ with $k = 4$.

F^2_{6,σ_0}: Formula (51) reduces the computation of $B(F^2_{6,\sigma_0}, 4)$ to that of $B(F_{6,\rho}, 2)$, where $\rho = (6, 4, 2, 5, 3, 1)$. However, we cannot apply (50) since $\lambda(\rho) = 2$ and the condition on k does not hold. Instead, we note that all the U_i's are non-empty and we use the value $\delta_i = 1/4$ in all computations. It can be shown that this choice leads to a curve length which is at most 1 percent below the length of the maximal curve.

$F^2_{6,\sigma_{13}}$: As in the case of F_{10,σ_7}, the maximal curve is obtained when U_i is non-empty only for even i. A simple computation then yields $B(F^2_{6,\sigma_{13}}, 4) = \frac{6}{\sqrt{12}}B(F_1, 2)$.

F^2_{6,σ_2}: By aisle symmetry, this involves the same calculation as F^2_{6,σ_0}.

Online Appendix E: Notation

Notation	Meaning
n	Number of passengers.
m	Number of rows.
h	Number of passengers per row.
l	Distance between successive rows.
W	Aisle length occupied by a passenger (personal space).
D	Delay time parameter (for random boarding).
D_{oi}	Delay time in outside-in policy.
T	Total boarding time in the tasks with precedences model.
q	The (normalized) queue position coordinate.
r	The (normalized) row position coordinate.
(q, r)	The pair of coordinates representing a passenger.
p	A joint probability distribution on (q, r). It is derived from the airline boarding policy.
$\alpha(q, r)$	A computed quantity which measures the ability of the passenger at position (q, r) to block other passengers. Depends on p, and hence on the policy.
k	Congestion, ratio of total queue length to aisle length.
σ, ρ	Generic notation for permutations on boarding groups.
G_i	Groups of a boarding policy.
S_i	The sub-squares of the unit square which represent passengers from group $G_{\sigma(i)}$.
R_i	The sub-rectangles of the unit square which represent passengers from group $G_{\sigma(i)}$, in a multi-class policy.
U	The maximal curve. The critical path tends (roughly) to follow this curve.
U_i	The portion of the critical curve which lies in S_i (or R_i). May be empty.
ϕ	Generic notation for a curve
$L(\phi)$	The length of a curve ϕ, given in (5)
$B(F, k)$	The (normalized) estimated boarding time for a policy F, given congestion k
Λ	A partition of the passengers into non-uniform groups.
Δ_i	Size of the i'th group in a non-uniform partition is $\Delta_i n$.
λ_i	An element in the partition. The passengers with row coordinate between λ_{i-1} and λ_i form the i'th group of the non-uniform policy.

Table 3: List of notations

Notation	Meaning
F	Generic notation for a boarding policy.
F_v	The uniform (equal group sizes) back-to-front policy with v groups.
$F_{v,\sigma}$	A uniform policy with v groups, which are called in the order given by σ.
F_v^2	A uniform policy with two half-row classes and a back-to-front policy with v groups in each class.
$F_{v,\sigma}^2$	A uniform policy with two half-row classes and v groups in each class, called by the order given by σ.
F_v^c	A uniform policy with c classes of randomly chosen passengers, and v groups per class in back-to-front ordering.
$F_{\Lambda,\sigma}$	The policy which boards the groups defined by the partition Λ in the order given by σ.
F_Λ	The non-uniform, back-to-front policy with groups given by the partition Λ.
F^{oi}	The outside-in policy, which first boards window passengers, then middle and then aisle.
F_v^{oi}	The outside-in policy coupled with a back-to-front policy with v groups in each class.
$F_{\mathrm{nu},1/3}^{\mathrm{oi}}$	The outside-in policy coupled with a non-uniform back-to-front policy which has two groups of size $n/3$ and $2n/3$, respectively.

Table 4: List of policies

Online Appendix F: Variants on the boarding model

In our boarding process model and its asymptotic analysis we have made several assumptions:

- The delay D is the same for all passengers.
- The aisle length W occupied by each passenger is the same.
- The distance between rows l is the same between any two successive rows
- The number of passengers per row h is fixed.
- The airplane is full. (This assumption was implicit.)
- Passengers strictly obey the airline policy.

In this appendix we show how to adjust the functional given in (5) so that the asymptotic analysis still holds when we eliminate these assumptions.

In an airplane, h and l may vary between different compartments. For instance, the first class compartment has more leg room and less passengers per row than the travel class compartment. As a result, in general, we will have functions $l = l(r)$ and $h = h(r)$, where r in this case represents aisle length measured according to some normalized units. We normalize the function $l(r)$ so that $\int_0^1 \frac{h(r)}{l(r)} dr = 1$. It may also happen that the airplane is not full, leading to an occupancy function $0 \le \Gamma(r) \le 1$ which represents the occupancy percentage near row r. To accommodate varying leg room and passengers per row, we define the leg room rate $\zeta(r) = 1/l(r)$. The row difference between aisle length positions r and $r+dr$ will be given by $\zeta(r)drn$. Integrating, we obtain $\int_0^1 \zeta(r)dr = 1/E(h)$.

We may view $h(r)\zeta(r)\Gamma(r)$ as measuring the density of passengers per unit aisle length. Consequently, we define the adjusted density function $\mu(q,r) = \chi p(q,r)h(r)\zeta(r)\Gamma(r)$, where χ is chosen so that $\int \mu(q,r)dqdr = 1$. Let W denote an aisle length distribution rather than a constant value. We assume that W is also measured with respect to the aisle length measure r, that is, n people will occupy roughly $E(W)$ units of aisle length, where $E(W)$ is the average of W. Put $\eta(q,r) = \int_r^1 \mu(q,z)dz$. We still assume that D is a constant. Consider the set of curves ϕ satisfying

$$\phi' + E(W)\eta \geq 0. \tag{52}$$

Consider the functional

$$L(\phi) = 2D \int_0^1 \sqrt{\mu(q,r)(\phi' + E(W)\eta)}dq, \tag{53}$$

and let

$$B = Max_{\phi} L(\phi), \tag{54}$$

where the maximum (which exists) is taken over all legitimate curves. Then $B\sqrt{n}$ is the asymptotic boarding time.

A more complicated problem arises if we wish to consider D as a delay distribution, rather than a constant. When D is a random variable, the boarding time is given by the maximal sum of delays in a precedence chain. When $k = 0$ and p is uniform, this leads to the problem of finding the highest weight increasing subsequence in a random permutation, where the weights are distributed according to D. By a theorem of Hammersley, the highest weight increasing subsequence will have w.h.p a weight asymptotic to $c_D\sqrt{n}$, where c_D is a constant depending on D. It is easy to show that $c_D \geq E(D)$, but strict inequality may hold. We can then apply the procedure above to obtain the functional

$$L(\phi) = \int_0^1 2c_D \sqrt{\mu(q,r)(\phi' + E(W)\eta)}dq. \tag{55}$$

Note that the formula is local, in the sense that it holds even when D depends (continuously) on q and r. Such dependence may occur since passengers near the end of the queue are likely to experience longer delays due to overhead bin and seat occupancy.

Passengers who do not adhere to the airline policy can be modeled by changing the density $p(q,r)$. In the case of announcement policies, this means that the density function $p(q,r)$ will not vanish outside the squares S_i (or rectangles R_i). Instead, it will have a positive value which depends on the percentage of policy defectors and possibly on the size of the infringement. Such changes do not affect the form of the functional. However, they tend to complicate the explicit calculations.

If an airplane has more than one aisle, as most large airplanes do, then we assume that passengers are directed to the aisle closest to their seat. This procedure is usually performed in person by a flight attendant located at the entrance of the airplane. Assuming that passengers are directed to the correct aisle, the airplane boarding problem will decompose into several, single aisle boarding problems, which are run in parallel. The main term asymptotics of the boarding time calculations will not change. We note, however, that directing of the passengers into the correct aisle may become the bottleneck for boarding time. This issue needs to be examined in field trials.