Stochastic Transit Equilibrium[☆]

Cristián E. Cortés[a,*], Pedro Jara-Moroni[b], Eduardo Moreno[c]

[a] *Department of Transport Engineering, Universidad de Chile*
[b] *Department of Economics, Universidad de Santiago de Chile*
[c] *Faculty of Engineering and Sciences, Universidad Adolfo Ibañez*

Abstract

We present a transit equilibrium model in which boarding decisions are stochastic. The model incorporates congestion, reflected in higher waiting times at bus stops and increasing in-vehicle travel time. The stochastic behavior of passengers is introduced through a probability for passengers to choose boarding a specific bus of a certain service. The modeling approach generates a stochastic common line problem, in which every line has a chance to be chosen by each passenger. The formulation is a generalization of deterministic transit assignment models where passengers are assumed to travel according to shortest hyperpaths. We prove existence of equilibrium in the simplified case of parallel lines (Stochastic Common-line Problem) and provide a formulation for a more general network problem (Stochastic Transit Equilibrium). The resulting waiting time and network load expressions are validated through simulation.

Keywords: Transit Equilibrium, Stochastic Models, Hyperpaths, Congested Networks, Simulation

[☆] The authors thank the financial support of Fondecyt Chile Grant 1100239, the Millennium Institute "Complex Engineering Systems" (ICM: P-05-004-F, CONICYT: 522 FBO16) and Anillo Grant ACT88.

[*] Corresponding author

Email addresses: ccortes@ing.uchile.cl (Cristián E. Cortés), pedro.jara@usach.cl (Pedro Jara-Moroni), eduardo.moreno@uai.cl (Eduardo Moreno)

1. Introduction

Public transport assignment models have been formulated to properly represent the way in which the passengers of a transit system utilize the available supply (in terms of infrastructure, line frequencies and other predefined operational rules) for travelling through the transit network from different origins to different destinations. This problem has been subject of many studies during the last decades, mainly oriented to systems of buses; most proposed models have been designed to properly represent the passenger behaviour when moving within the transit network, with the objective of predicting equilibrium conditions under hypothetical scenarios, with the final goal of studying the potential benefits of high impact public transport projects in large cities and metropolitan areas. For strategical purposes, these models have been incorporated in more general frameworks, mostly with the objective of reaching global traffic equilibrium conditions in cases where multiple transport modes interact.

In terms of passenger behaviour, the recent literature has been oriented to model passenger preferences assuming that they use path selection strategies to reach their destinations. Originally, Spiess and Florian (1989) define a strategy as a set of rules that, when applied, allows the passenger to reach his(her) destination, and the decisions are made at each node where boarding is allowed. A properly defined strategy includes the choice of sets of attractive lines at bus stops (also called common lines) as described by Chriqui and Robillard (1975) and Spiess and Florian (1989). The notion of strategy implies that passengers have good knowledge of the network structure and conditions; therefore they are able to identify and utilize effective strategies (Bouzaiene-Ayari et al., 2001). The problem of minimizing the expected cost - that should include at least in-vehicle travel, access and waiting times for passengers - can be then modelled as a user equilibrium problem on the hyperpath space, concept introduced in graph-theory language by Nguyen and Pallottino (1989) and then applied to a transit assignment problem (Nguyen and Pallottino, 1988). An hyperpath is basically an acyclic subgraph connecting a single origin-destination, paired with a given vector of real arc values. Nguyen and Pallottino (1989) show that the notion of hyperpath can be derived from a relaxation of the definition of a path. This description from graph theory permits converting the passenger assignment problem into a standard equilibrium problem on a private car network.

The first studies did not consider a relevant issue in passenger transit assignment, namely the congestion produced at bus stops when the capacity of transport is not enough to serve the demand for that service. The original models without congestion were reasonable under low passenger demand conditions at bus stops (Nguyen and Pallottino, 1988, 1989; Spiess, 1984; Spiess and Florian, 1989), and in those cases, the result of the assignment models is equivalent to finding the equilibrium of the transit system. If that is not the case (for example in many Latin American capitals during peak hours), the reduced capacity is reflected in higher waiting times for passengers as they can not always board the next bus arriving to the stop. The issue of capacity was first treated by Gendreau (1984) who generated much complex formulations as the waiting process was based on a bulk queue model, making impossible to properly formulate the equilibrium under congestion at bus stops. De Cea and Fernández (1993) developed an alternative model assuming that passengers travel through a sequence of successive intermediate nodes, allowing choices among multiple lines at a given stop only in they all share the next stop to be served (simplified version of an hyperpath-based formulation). The authors were able to heuristically incorporate congestion at bus stops, through a model built on an augmented graph very expensive from a computational standpoint, and where the obtained flows can exceed the capacity of the vehicles as the functional form used to represent congestion was not asymptotic on capacity. In addition, in this model the common lines between intermediate nodes are computed heuristically and therefore there is no guarantee of reaching equilibrium conditions.

Wu et al. (1994) studied a congested network assignment model with passengers travelling according to shortest hyperpaths. Travel times as well as waiting times are considered to be flow dependent, but the passenger assignment is based on the nominal frequencies of the lines. Bouzaiene-Ayari et al.

(2001) point out the difficulty of modelling bus stops, and extend the Wu et al. (1994) model to study existence, uniqueness and equilibria in case where the flow distribution is done proportionally to the inverse of the waiting time of each line. Their proposal does not permit the use of congestion functions borrowed from queuing theory (as under congestion waiting times go to infinity) and also assume travel-time functions to be strongly monotone, which prevents the model from being used in the case of constant travel times. Cominetti and Correa (2001) analyse an hyperpath-based equilibrium model for passenger assignment in general transit networks including explicitly the congestion effects at bus stops over the passenger' choices. Congestion is treated by means of a bulk queue model at the stops. The authors provide a complete characterization of the set of equilibria in the common-line setting, including the conditions for existence and uniqueness. They show that over certain ranges, an increase of flow does not affect the system performance in terms of transit times. The authors study a general equilibrium model supporting multiple origins and destinations, overlapping bus lines, as well as transfers at intermediate nodes on a given trip; the authors model this general case through a dynamic programming approach for representing a common-line scheme including congestion effects, and are able to establish the existence of a network equilibrium. Cepeda et al. (2006) extend the formulation by Cominetti and Correa (2001) obtaining a new characterization of the equilibria in the context of a congested transit networks with capacity constraints at bus stops; by using this approach it is possible to formulate an optimization problem in terms of a computable gap function that vanishes if the solution reaches equilibrium. The method leads to an algorithm that uses the method of successive averages (MSA) from where they are able to find equilibrium conditions on large-scale networks with congestion. Schmcker et al. (2008) develop a capacity constrained transit assignment model in a dynamic fashion, allowing passengers not able to baord a vehicle in a previous period, to be transferred to the next interval. The common line problem is considered and the search for the shortest hyperpath is influenced by a fail-to-board probability introduced by the potential overcrowding (for example during peak periods) at certain intervals. The dynamic approach adds a priority rule in the network loading process not able to properly consider in a static assignment.

The above discussed models add a relevant feature in transit assignment (and in some cases in transit equilibrium) modelling, which is the inclusion of congestion at stops (or stations) as part of the proper representation of passenger behaviour; the congestion is related to the impact on the system performance due to capacity constraints associated with the finite size of vehicles. This phenomenon precludes some passengers to board the vehicles, in all these cases due to a hard constraint. However, it could be the case where passengers are observed not to board a bus of certain line even though the bus has capacity available. We claim that there are other external conditions that could modify the passenger behaviour in transit assignment not related to capacity constraints. The work by Nguyen et al. (1998) shows a stochastic assignment model based on the hyperpath framework for transit networks. The stochasticity added in this case is through a Logit assignment structure at the hyperpath decision of the passengers at the boarding nodes. The authors are able to model stochasticity in the context of transit assignment, although they neither consider more general conditions that could lead to a global transit equilibrium, nor add capacity constraints at stops.

The goal of this paper is to add the stochastic effect into boarding decisions at bus stops, in the context of a transit equilibrium model with congestion at bus stops, reflected in potential higher waiting times due to overcrowding of the system. The model is an extension of the proposal of Cominetti and Correa (2001) and Cepeda et al. (2006), where passengers are assumed to to travel according to shortest hyperpaths. Travel times are not necessarily monotone and congestion affects both the waiting times and the flow distribution. The stochastic behaviour of passengers is introduced through a distribution of probabilities for passengers to board a specific bus of certain service that can be characterized by its observed frequency at that stop and its travel time to the next stop. The modelling approach generates a stochastic common line problem, in which every line has a chance to be chosen by each passenger, even if the service quality offered by the line is quite poor. The formulation also incorporates capacity constraints due to overcrowding at stops in the same way as Cepeda et al. (2006) propose. The formulation is a generalization of the hyperpath model for the deterministic case, which can be recovered by properly setting the probability of choosing the available lines in one or zero depending on the service provided by the line.

We prove existence of equilibrium in the simplified case of parallel lines (stochastic common line problem) to show the consistency of the proposed model, extending later the formulation to a more general network problem. Under this stochastic formulation for the more general model, the recursive expressions for the time-to-destination functions can be analytically found together with the line flows at equilibrium, by solving a set of simultaneous stochastic common-line problems (one for each origin-destination pair), coupled by flow conservation constraints; note the difference with the deterministic models with congestion (Cominetti and Correa, 2001; Cepeda et al., 2006), in which finding the equilibrium requires solving a set of generalized Bellman equations.

Thus, a stochastic model in passenger behaviour allows us to provide a more realistic representation of the decisions made by passengers in a context of equilibrium under bus capacity congestion based on minimum hyperpaths. Linked to that, the implementation of the model is less cumbersome and allows modelling other line penalties different from waiting and travel times through a proper expression of the choice probabilities. As it seems quite difficult to write analytical expressions for the expected waiting times at stops when including stochastic behaviour in the passenger assignment model with congestion, we decide to validate our proposed stochastic formulations through simulation, analysing different cases in terms of demand, line probabilities and arrival rates of passengers.

The paper is organized as follows. In Section 2 the transit equilibrium model for the simple case of one origin-destination is presented, which is then extended to a general network formulation at the end of the section. Next, in Section 3, a queuing model for a single stop and multiple lines is formulated and later validated through simulation. In Section 4, final remarks and further developments

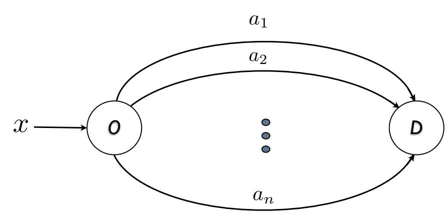

Figure 1: A simple network: an origin destination pair connected by parallel links.

are highlighted.

2. Transit model

2.1. Stochastic Common Lines.

We propose a stochastic common line approach in which perception of travel time is random across passengers.

Consider the network depicted in Figure 1 consisting of an origin O and a destination D node, connected by a finite set of arcs or links A. Each arc a represents a bus line that serves the origin/destination pair (OD pair). Each line $a \in A$ is characterized by two elements. The first one is a constant in vehicle travel time $t_a \in \mathbb{R}_+$. The second one is the frequency of service of the line f_a. Since large flows and limited capacity of buses may prevent passengers form boarding a bus, congestion at the bus stop increases their waiting times. To model this situation we will assume that the frequency of service of each line is modeled by a strictly decreasing and smooth *effective frequency* function of the flow on bus line a, v_a, $f_a : [0, \bar{v}_a[\rightarrow]0, +\infty[$ that vanishes at $\bar{v}_a$. This is $f_a \rightarrow 0$ when $v_a \rightarrow \bar{v}_a$ (Cominetti and Correa, 2001).

A stochastic model provides the probability p_a of a passenger wishing to board a bus at the bus stop, given that a bus of line a has arrived at the bus stop:

$$p_a = \mathbb{P}(\text{wishes to board bus}|\text{bus of line } a \text{ is at stop}) \,.$$

Each passenger that wishes to travel from O to D, compares the travel time on the current bus, with the expected travel time of waiting for the next bus:

$$\begin{array}{ll} \text{if passenger boards bus} & \rightsquigarrow \text{travel time } t_a; \\ \text{if passenger does not board bus} & \rightsquigarrow \text{travel time } T. \end{array}$$

Assuming that the bus arrival process is completely renewed when not boarding, expected travel time can be calculated as:

$$T = \frac{1}{\sum_a f_a(v_a)} + \sum_a \frac{f_a(v_a)}{\sum_{a'} f_{a'}(v_{a'})} \left[p_a t_a + (1 - p_a)T\right]. \tag{1}$$

The first term is the standard expression for expected waiting time while the second term is the expected travel time, that takes into account the possibility of not boarding the bus at the bus stop, in which case the passenger will spend an extra expected time T.

Clearing T in (1) we get an expression $T(v, p)$.

$$\begin{aligned} T(v,p) &= \frac{1 + \sum_a f_a(v_a)\, p_a t_a}{\sum_{a'} f_{a'}(v_{a'}) - \sum_a f_a(v_a)\,(1 - p_a)} = \frac{1 + \sum_a f_a(v_a)\, p_a t_a}{\sum_a f_a(v_a)\, p_a} \\ T(v,p) &= \frac{1}{\sum_a f_a(v_a)\, p_a} + \sum_a t_a \frac{f_a(v_a)\, p_a}{\sum_{a'} f_{a'}(v_{a'})\, p_{a'}} \end{aligned} \tag{2}$$

Considering (2), we can re-interpret equation (1). The expression $\frac{1}{\sum_a f_a p_a}$ can be interpreted as an expected waiting time that takes into account the stochastic model of boarding, while expression $\frac{f_a p_a}{\sum_{a'} f_{a'} p_{a'}}$ is the probability of boarding bus line a. In Section 3 we study the validity of these two expressions.

Consider a flow $x > 0$ of passengers that wish to travel from O to D. Total flow splits among all possible bus lines so that $x = \sum_{a \in A} v_a$. Since the probability with which passengers board bus line a is $\frac{f_a p_a}{\sum_{a'} f_{a'} p_{a'}}$, bus load at a bus stop can then be calculated by the system of equations:

$$v_a = x \frac{f_a(v_a)\, p_a}{\sum_{a'} f_{a'}(v_{a'})\, p_{a'}} \qquad a \in A. \tag{3}$$

Expression (2) along with the system (3), provide a set of equations on $v = (v_a)_{a \in A}$ and $p = (p_a)_{a \in A}$. A solution to this extended system is a transit equilibrium of this simple network.

Deterministic Case. If decisions are deterministic, then

$$p_a = \begin{cases} 0 & \text{if } t_a > T \\ 1 & \text{if } t_a < T \end{cases}$$

This leads to the common lines paradigm. Indeed, solving $\min_p T(v,p)$ we obtain the solution to the common lines problem (see Chriqui and Robillard (1975); for a treatment with congestion see De Cea and Fernández (1993)). This particular case of common lines under congestion is rigourously studied in Cominetti and Correa (2001) and Cepeda et al. (2006).

Non Deterministic Case. If decisions are stochastic, then each passenger has probability p_a of wishing to board a bus of line a when this bus is at the bus stop. This probability is given by a stochastic model and depends on the expected travel time. We will assume that p_a is a strictly decreasing continuous function of the difference between travel time in line a and expected travel time T, $\varphi_a : \mathbb{R} \to \,]0,1[$:

$$p_a \equiv \varphi_a(t_a - T)\,. \tag{4}$$

Under this formulation, it is no longer possible to obtain T as a function of (v,p) as in (2). We have to state the equilibrium conditions in terms of the variable τ that represents equilibrium expected travel time from O to D. A second set of equations comes from the network load (3). We may now state a definition of equilibrium for our model.

Definition 1. *A* Stochastic Common-line Equilibrium *is a pair* $(v^*, \tau^*) \in \prod_{a\in A} [0, \bar{v}_a[\times \mathbb{R}_+$, *such that*

$$\begin{aligned} &\tau^* = T(v^*, p)\,; \\ &v_a^* = x\frac{p_a f_a(v_a^*)}{\sum_{a'} f_{a'}(v_{a'}^*)\, p_{a'}} \qquad \forall\ a \in A; \end{aligned}$$

and

$$p_a = \varphi_a(t_a - \tau^*) \qquad \forall\ a \in A.$$

2.2. Existence of Stochastic Common-line Equilibrium

We now prove existence of equilibrium for the simple network of the previous section.

Proposition 1. *Consider a flow $x > 0$ of passengers that wish to travel from O to D. If $\sum_{a\in A} \bar{v}_a > x$ and for all $a \in A$, the functions φ_a are differentiable for all $s > 0$ and satisfy:*

$$s\varphi_a'(s) + \varphi_a(s) > 0 \qquad \forall s > 0, \tag{5}$$

then there exists a Stochastic Common-line Equilibrium in the network with one OD pair and n parallel links.

Proof. Condition $\tau^* = T(v^*, p)$ comes originally from equation (1). If in (1) we replace, for each $a \in A$ the values p_a by the functions φ_a we may obtain:

$$0 = 1 + \sum_a f_a(v_a) \cdot (t_a - T) \cdot \varphi_a(t_a - T)\,; \tag{6}$$

and so equation (6) relates T to v. For a given v the right hand side of (6) tends to $-\infty$ when $T \to +\infty$ and it is positive in $T = 0$. Therefore, since for all a, φ_a is continuous, equation (6) has a solution in T. Condition (5) implies that the right hand side of (6) is strictly increasing as a function of T and so this solution is unique. Let us denote it $T(v)$.[1]

Coupling (4) and (3) we obtain:

$$v_a = x\frac{f_a(v_a)\,\varphi_a(t_a - T(v))}{\sum_{a'} f_{a'}(v_{a'})\,\varphi_{a'}(t_{a'} - T(v))} \qquad \forall a \in A. \tag{7}$$

Let us consider the right hand side of the previous expression as a function of v. Clearly it is well defined for $v \in \prod_{a\in A} [0, \bar{v}_a[$. Condition (5) implies that T is a continuous function[2] and so the right hand side of (7) is as well a

[1] Note that condition (5) holds for all $s \leq 0$. It states that the probability φ_a must go to zero *faster* than $-\frac{\varphi_a(s)}{s}$.

[2] The Implicit Function Theorem holds (See for instance Simon and Blume, 1994, Theorem 15.2, page 341).

continuous function. Moreover, this function may be extended continuously to the set $\mathcal{V} \equiv \left\{ v \in \mathbb{R}_+^{|A|} \ : \ \sum_{a \in A} v_a < \sum_{a \in A} \bar{v}_a \right\}$.

We define then, for each $a \in A$, the function $V_a : \mathcal{V} \to [0, x]$ as:

$$V_a(v) := \begin{cases} x \frac{f_a(v_a)\varphi_a(t_a - T(v))}{\sum\limits_{a' \text{ s.t. } v_{a'} < \bar{v}_{a'}} f_{a'}(v_{a'})\varphi_{a'}(t_{a'} - T(v))} & \text{if } v_a < \bar{v}_a \\ 0 & \text{if } v_a \geq \bar{v}_a. \end{cases}$$

Note that if $v_{a'} \geq \bar{v}_{a'} \ \forall \ a' \neq a$, then $V_a(v) = x$. Furthermore, for $v \in \mathcal{V}$

$$\sum_{a \in A} V_a(v) = x.$$

We may define then the function $V : \mathcal{V}(x) \to \mathcal{V}(x)$ as:

$$V(v) := \prod_{a \in A} V_a(v)$$

with $\mathcal{V}(x) := \left\{ v \in \mathbb{R}_+^{|A|} \ : \ \sum_{a \in A} v_a = x \right\}$.

Since $\sum_{a \in A} \bar{v}_a > x$, V is a well defined continuous function, from $\mathcal{V}(x)$ in itself. The set $\mathcal{V}(x)$ is compact, convex and non empty and so V has fixed point v^*. Defining $\tau^* := T(v^*)$ it is direct to see that (v^*, τ^*) is a stochastic common-line equilibrium.

□

2.3. General formulation

We now formulate the stochastic equilibrium model for general transit networks. Network structure and notation follow Cepeda et al. (2006), who in turn consider transit networks as built in Spiess and Florian (1989).

The formulation is developed on a general directed graph $G = (N, A)$. We denote by i_a and j_a respectively the tail and head nodes of a link $a \in A$, and we let $A_i^+ = \{a \in A : i_a = i\}$ and $A_i^- = \{a \in A : j_a = i\}$ be the sets of arcs leaving and entering node $i \in N$.

The set of destinations is denoted $D \subset N$, and for each $d \in D$ and every node $i \neq d$ a fixed demand $g_i^d \geq 0$ is given.[3] To keep the model tractable we need to specify arc-destination flows. The set $\mathcal{V} := \mathbb{R}_+^{|A|\times|D|}$ denotes the space of arc-destination flow vectors v with nonnegative entries $v_{ad} \geq 0$, while $\mathcal{V}_0$ is the set of feasible flows $v \in \mathcal{V}$ such that $v_{ad} = 0$ for all $a \in A_d^+$ (i.e. no flow with destination d exits from d) and satisfying the flow conservation constraints

$$g_i^d + \sum_{a \in A_i^-} v_{ad} = \sum_{a \in A_i^+} v_{ad} \qquad \forall i \neq d.$$

The formulation in this section differs from the previous one in that now we will allow in vehicle travel time t_a and the effective frequency functions to depend on the complete vector of link flows v. We introduce this modification because when we study passenger assignment and stochastic transit equilibrium in more general networks this dependence is unavoidable. In vehicle travel time in a specific link of a transit network is indeed affected by the flow of passengers that board the bus at the end of the link.[4] Similarly, waiting times do not only depend on the boarding flows and operational characteristics of the lines but also on the on-board flows which consume part of the line capacity.

To be precise, we assume that each link $a \in A$ is characterized by a continuous travel time function $t_a : \mathcal{V} \to [0, \bar{t}_a[$, where $\bar{t}_a$ is a finite upper bound, and the effective frequency function $f_a : \mathcal{V} \to [0, +\infty]$ which is either identically $+\infty$ or everywhere finite, in which case, for each $d \in D$ we assume that $f_a \to 0$ when $v_{ad} \to +\infty$ with $f_a(v)$ strictly decreasing with respect to v_{ad} when strictly positive.

The intuitive idea behind the notion of a stochastic transit equilibrium follows directly from Cominetti and Correa (2001) and Cepeda et al. (2006). Consider a passenger heading towards destination d and reaching an intermediate node i in his trip (see Figure 2). To exit from i he can use the arcs $a \in A_i^+$

[3] Usually the demands g_i^d are strictly positive only at nodes i corresponding to stop-nodes, that is to say the bus stops where users wait for service, but no restriction is imposed.

[4] Of course, in the simple network framework this phenomenon is not present since no passengers board buses at D.

to reach the next node j_a. By taking the arc travel times $t_a(v)$ and the transit times $\tau_{j_a d}$ from j_a to d as fixed, the decision faced at node i is a common-line problem with travel times $t_a(v) + \tau_{j_a d}$ and effective frequencies corresponding to the services operating on the arcs $a \in A_i^+$. The solution of this stochastic common-line problem determines the transit time τ_{id} from i to d, which can then be used recursively to solve the upstream nodes. Time to destination $\tau_{id}(v)$ from each node i to destination d is obtained from the stochastic common-line problem from i to d characterized in Definition 1 with $t_a \to t_a(v) + \tau_{j_a d}$ if $i \neq d$ and $\tau_{dd} = 0$.[5]

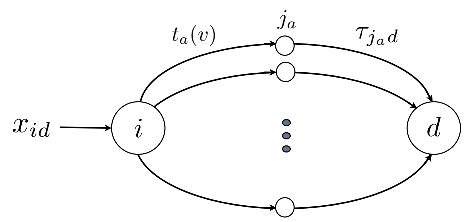

Figure 2: The Common-line problem and the Hyperpath concept.

All variables $\tau_{j_a d}$ and v_{ad} must be determined at the same time, so the stochastic transit equilibrium is formulated as a set of simultaneous stochastic common-line problems (one for each *id* pair), coupled by flow conservation constraints. We define for each $v \in \mathcal{V}$ the *flow entering node i with destination d* by:

$$x_{id}(v) := g_i^d + \sum_{a \in A_i^-} v_{ad}.$$

Definition 2. *A pair of feasible flow vector and expected travel times $(v^*, \tau^*) \in \prod_{a \in A} [0, \bar{v}_a[\times \mathbb{R}_+^{|N| \times |D|}$ is a Stochastic Transit Equilibrium if for all $d \in D$ and*

[5]Note that $t_a(v)$ does not depend on the flow on arc a and so can be considered constant in the stochastic common line problems where it participates.

$i \in N$, with $i \neq d$ we have:

$$\tau^*_{id} = \frac{1 + \sum_{a \in A_i^+} p_a^d f_a(v^*) \Big(t_a(v^*) + \tau^*_{j_a d} \Big)}{\sum_{a \in A_i^+} f_a(v^*)\, p_a^d} \qquad ;$$

$$v^*_{ad} = x_{id}(v^*)\, \frac{f_a(v^*)\, p_a^d}{\sum_{a \in A_i^+} f_a(v^*)\, p_a^d} \quad , \qquad \forall a \in A_i^+;$$

$$p_a^d = \varphi\big(t_a(v^*) + \tau^*_{j_a d} - \tau^*_{id}\big) \quad , \qquad \forall a \in A_i^+.$$

The conditions in a Stochastic Transit Equilibrium are a direct adaptation from the deterministic ones that characterize a Transit Network Equilibrium (Cominetti and Correa, 2001; Cepeda et al., 2006). A significant difference that arises from the incorporation of a stochastic model of boarding, is that we withdraw strategies as a modeling tool since at every bus stop every line has a positive boarding probability. Consequently expected travel time, τ, is obtained directly by a (recursive) functional form instead of a dynamic programming problem.

3. Queuing model

In this section, we focus the public transport system analysis on the operation of one isolated bus stop served by L bus lines. The goal is to provide a queue-theoretic framework to support the stochastic assignment model with congestion that we are proposing in Definition 1, justifying the formulae

$$\text{Expected waiting time:} \qquad \frac{1}{\sum_a f_a(v_a) p_a} \tag{8}$$

$$\text{Probability of boarding bus line } a\text{:} \qquad \frac{f_a(v_a) p_a}{\sum_{a'} f_{a'}(v_{a'}) p_{a'}} \tag{9}$$

3.1. Uncapacited bus lines

First of all, we prove that equations (8) and (9) hold under no capacity constraints on the buses. First, we study the case of a single bus line with Poisson arrivals of rate μ and a boarding probability p. Let X_i be the time between arrivals of two consecutive buses. Therefore, the waiting time of a passenger is equal to $\sum_{i=1}^{k} X_i$ with probability $(1-\rho)^{k-1}\rho$. Since X_i follows an exponential distribution with parameter μ, the expected waiting time is given by

$$W = \sum_{k=0}^{\infty} (k+1) \frac{1}{\mu} (1-p)^k p = \frac{1}{p\mu}$$

Suppose now the case of L parallel lines with rates μ_a and boarding probabilities p_a. We can model the arrival of buses as one Poisson process of rate $\sum_{a'} \mu_{a'}$. Hence, the expected waiting time W of passenger is given by

$$W = \frac{1}{\bar{p}\left(\sum_{a'} \mu'_a\right)}$$

where $\bar{p}$ is the probability of boarding the bus at the bus stop, which is given by

$$\bar{p} = \sum_a p_a \left(\frac{\mu_a}{\sum_{a'} \mu_{a'}} \right)$$

replacing this equation in previous formula, we obtain

$$W = \frac{1}{\sum_a p_a \mu_a}$$

On the other hand, the probability of a passenger to finally board bus line a is given by

$$\delta_a = \sum_{i=0}^{\infty} p_a \left(\frac{\mu_a}{\sum_{a'} \mu_{a'}} \right) (1-\bar{p})^i = \frac{p_a}{\bar{p}} \left(\frac{\mu_a}{\sum_{a'} \mu_{a'}} \right) = \frac{p_a \mu_a}{\sum_{a'} p_{a'} \mu_{a'}}$$

3.2. Capacited bus lines

We now study the case where buses arrive with a limited capacity. The objective is to obtain analytical expressions to compute the effective frequency function as well as the expected waiting times in the case of including stochastic behaviour in the passenger assignment model. The model is an extension of the one shown in Appendix A of Cominetti and Correa (2001). In this case, passengers arrive according to a Poisson process of rate v, while buses arrive as a Poisson process as well, with rate μ for the single line serving the stop, with random available capacity C, where $\mathbb{P}(C = j) = q_j$, $j = 0, \ldots, K$. In Cominetti and Correa (2001) if a bus arrives with available capacity larger than the queue length, the latter reduces to zero. The difference with the Cominetti and Correa

(2001) model is that in this stochastic version, some passengers could eventually decide not to get on a certain bus and wait for the next one, even if that bus had available capacity to accommodate those passengers.

As before, assuming that p is the probability with which a certain passenger boards a random bus at the bus stop, then the queue length is a continuous time Markov chain with transition rates

$$
\begin{aligned}
\theta_{k,k+1} &= v, && k \geq 0; \\
\theta_{k,k-j} &= \mu \Big[\mathbb{P}(C > j)\, \mathbb{P}(\text{board } j | \text{there are } k) && k \geq 1 \text{ and} \\
&\quad + \mathbb{P}(C = j)\, \mathbb{P}(\text{board} \geq j | \text{there are } k) \Big], && 0 \leq j \leq \min\{k, K\}; \\
\theta_{k,0} &= \mu p^k \sum_{j=k}^{K} q_j, && 1 \leq k \leq K;
\end{aligned}
$$

where the probabilities $\mathbb{P}(\text{board } j | \text{there are } k)$ and $\mathbb{P}(\text{board} \geq j | \text{there are } k)$ can be computed as follows

$$\mathbb{P}(\text{board } j | \text{there are } k) = \binom{k}{j} p^j (1-p)^{k-j} \tag{10}$$

$$\mathbb{P}(\text{board} \geq j | \text{there are } k) = \sum_{l=j}^{k} \binom{k}{l} p^l (1-p)^{k-l} \tag{11}$$

The stationary distribution $\Pi = \{\pi_k\}_{k \geq 0}$ is characterized by the balance equations obtained from the solution of the system $P^T \Pi = \Pi$, where P is the matrix of transition probabilities of the queue length. These probabilities are equal to the transition rates divided by $(v + \mu)$. We solve then the system $\Theta^T \Pi = \Pi (v + \mu)$, considering $\theta_{0,0} = \mu$. In Appendix A the analytical calculations of this queuing model considering the stochastic case are shown in detail, obtaining the following system:

$$
\begin{aligned}
v\pi_0 &= \mu \sum_{l=1}^{K} q_l \sum_{k=1}^{l} \pi_k p^k; \\
(v+\mu)\,\pi_k &= \pi_{k-1} v + \mu \sum_{l=0}^{K} q_l \Bigg[\sum_{j=0}^{l} \pi_{k+j} \mathbb{P}\big(\text{board } j \big| \text{there are } k+j\big) \\
&\qquad + \pi_{k+l} \mathbb{P}\big(\text{board} > l \big| \text{there are } k+l\big) \Bigg], \qquad \forall k \geq 1.
\end{aligned}
\tag{12}
$$

The problem with the obtained balance equations is the fact that the coefficients associated with the π_k depend on the equation being considered. In other words, equation (12) contains coefficients that depend on k and that makes impossible to solve such a system analytically. Writing explicitly the probability expressions (10) and (11) we get

$$
\begin{aligned}
0 &= \left(-v - \mu\left(1-q_0\right)\left(1-(1-p)^k\right)\right)\pi_k + v\pi_{k-1} \\
&+ \mu \sum_{l=1}^{K} q_l \left[\sum_{j=1}^{l} \binom{k+j}{j} p^j \left(1-p\right)^k + \pi_{k+l} \sum_{m=l+1}^{k+l} \binom{k+l}{m} p^m \left(1-p\right)^{k+l-m} \right].
\end{aligned}
$$

As the previous expression could not be solved analytically through traditional stochastic processes and queuing theory techniques, even for the simplest case of one line and one bus stop, we decide to test the validity of our stochastic formulation through simulation, analysing different cases in terms of demand, line probabilities and arrival rates. The details of these experiments and the conclusions obtained from them regarding the consistency of the proposed stochastic transit equilibrium model are highlighted in the next subsection.

3.3. Simulation experiments

First, we simulate the arrival of passengers at a rate v to a bus stop served by L bus lines with arrival rates μ_a and boarding probabilities p_a. Each bus arrives to the bus stop with a random available capacity following an uniform distribution between 0 and Q. At the arrival of a bus of line i, each passenger in the queue decides to get on the bus or wait for the next one (according to probability p_a). If the number of passengers willing to board the bus is greater than its available capacity, we select a random subset of these passengers that board the bus. Using this simulation, we compute the empirical waiting time $\hat{W}$ and the fraction of passengers $\hat{\delta}_a$ that board line a, for each $a = 1 \dots L$.

In order to validate equations (8) and (9), we need to estimate the term $p_a f_a(v)$ for each bus line a under capacity constraints. In order to do that, a second simulation is performed simulating each line individually. In these cases, we assume that passengers arrives at a rate proportional to the fraction of passengers that board this line in the first simulation $v_a = \hat{\delta}_a v$ and we repeat the same steps of the first simulation, but this time with a single bus line a, for each $a = 1 \dots L$. On each line, we estimate the value of $p_a f_a$ as $\frac{1}{\bar{W}_a}$, where $\bar{W}_a$ is the average waiting time of passengers under the single-line simulation.

In Table 1 we show the results obtained in this simulation, for the case of two lines ($L = 2$), with different nominal frequencies μ_1, μ_2; probabilities p_1, p_2; and maximum capacity Q. In each simulation, we assume a passenger arrival rate of $v = 100$, and we simulate 1 million events (including passengers and bus arrivals). In column "Sim" we show the resulting waiting time $\hat{W}$ and fraction of passengers boarding the first line $\hat{\pi}_1$, obtained by the first simulation. In column "Est" we show the estimated parameters obtained by applying equations (8) and (9), computed using the average waiting times $\bar{W}_1$, $\bar{W}_2$ from the single-line simulation. Finally, in column "Gap" we compute the relative difference between columns "Sim" and "Est".

As we can see, the simulated waiting times and probabilities are similar to the estimated values obtained from single-line simulations, with average differences less than 2%. In particular, we can see that the deterministic case ($p_1 = p_2 = 1$) under high congestion obtains gaps even greater than the average gap of the stochastic cases. The higher gaps are obtained on the simulations under heavy congestion, asymmetric boarding probabilities and asymmetric nominal frequencies. This systematic bias needs further investigation, but it appears to come from the congestion, and not from the stochastic behaviour of passengers. In fact, with higher capacities these differences exceptionally exceed the 1%.

	$[p_1, p_2]$		$Q = 42$ Sim	Est	gap	$Q = 80$ Sim	Est	gap	$Q = 160$ Sim	Est	gap
$[\mu_1, \mu_2] = [12, 12]$	$[1, 1]$	W	0.046	0.046	-0.57%	0.044	0.044	-0.36%	0.042	0.042	0.30%
		δ_1	0.500	0.499	0.31%	0.501	0.498	0.70%	0.498	0.502	-0.81%
	$[1, 0.5]$	W	0.062	0.063	-0.60%	0.058	0.058	0.52%	0.057	0.057	0.11%
		δ_1	0.649	0.641	1.28%	0.659	0.658	0.15%	0.660	0.661	-0.07%
	$[0.8, 0.8]$	W	0.057	0.057	-0.24%	0.054	0.054	0.38%	0.053	0.053	-0.33%
		δ_1	0.501	0.500	0.17%	0.502	0.498	0.79%	0.501	0.501	0.11%
	$[0.8, 0.4]$	W	0.076	0.076	-0.25%	0.072	0.073	-1.02%	0.070	0.071	-0.23%
		δ_1	0.653	0.648	0.78%	0.658	0.658	0.05%	0.664	0.663	0.05%
	$[0.5, 1]$	W	0.062	0.062	-0.62%	0.058	0.059	-0.81%	0.057	0.057	0.23%
		δ_1	0.349	0.358	-2.70%	0.340	0.345	-1.55%	0.335	0.337	-0.77%
	$[0.5, 0.5]$	W	0.089	0.090	-0.67%	0.087	0.086	0.60%	0.084	0.084	-0.18%
		δ_1	0.497	0.500	-0.65%	0.501	0.500	0.20%	0.499	0.500	-0.27%
	$[0.4, 0.8]$	W	0.076	0.077	-0.90%	0.072	0.073	-0.87%	0.070	0.071	-0.81%
		δ_1	0.344	0.351	-2.12%	0.337	0.344	-1.95%	0.336	0.338	-0.42%
	$[0.4, 0.4]$	W	0.111	0.112	-0.23%	0.108	0.107	0.94%	0.105	0.105	0.07%
		δ_1	0.502	0.500	0.35%	0.499	0.500	-0.24%	0.501	0.502	-0.09%
$[\mu_1, \mu_2] = [12, 6]$	$[1, 1]$	W	0.065	0.065	-0.32%	0.059	0.060	-0.36%	0.057	0.057	-0.12%
		δ_1	0.664	0.665	-0.13%	0.665	0.666	-0.28%	0.668	0.667	0.17%
	$[1, 0.5]$	W	0.078	0.079	-1.74%	0.071	0.072	-1.19%	0.069	0.069	0.35%
		δ_1	0.785	0.776	1.13%	0.791	0.789	0.22%	0.796	0.796	0.03%
	$[0.8, 0.8]$	W	0.079	0.079	0.48%	0.074	0.073	0.12%	0.071	0.071	0.19%
		δ_1	0.668	0.667	0.19%	0.666	0.664	0.32%	0.665	0.664	0.14%
	$[0.8, 0.4]$	W	0.095	0.096	-1.33%	0.089	0.088	0.93%	0.085	0.086	-0.95%
		δ_1	0.785	0.780	0.66%	0.794	0.793	0.15%	0.795	0.795	0.00%
	$[0.5, 1]$	W	0.096	0.099	-3.02%	0.088	0.090	-1.30%	0.086	0.086	-0.25%
		δ_1	0.525	0.550	-4.73%	0.513	0.520	-1.32%	0.506	0.509	-0.60%
	$[0.5, 0.5]$	W	0.122	0.123	-0.50%	0.116	0.116	-0.32%	0.113	0.114	-0.30%
		δ_1	0.666	0.666	0.02%	0.666	0.667	-0.07%	0.665	0.666	-0.12%
	$[0.4, 0.8]$	W	0.118	0.120	-1.56%	0.111	0.111	-0.02%	0.107	0.107	-0.40%
		δ_1	0.523	0.540	-3.20%	0.512	0.516	-0.78%	0.505	0.507	-0.41%
	$[0.4, 0.4]$	W	0.152	0.153	-0.20%	0.146	0.145	0.99%	0.142	0.142	0.05%
		δ_1	0.665	0.668	-0.47%	0.668	0.666	0.26%	0.666	0.666	-0.02%
$[\mu_1, \mu_2] = [6, 3]$	$[1, 1]$	W	0.184	0.178	3.41%	0.133	0.133	-0.09%	0.120	0.120	0.00%
		δ_1	0.666	0.667	-0.11%	0.664	0.667	-0.47%	0.663	0.664	-0.11%
	$[1, 0.5]$	W	0.210	0.224	-6.64%	0.157	0.162	-3.08%	0.142	0.144	-1.14%
		δ_1	0.756	0.716	5.30%	0.785	0.772	1.69%	0.793	0.789	0.54%
	$[0.8, 0.8]$	W	0.206	0.205	0.40%	0.160	0.159	0.47%	0.148	0.148	-0.12%
		δ_1	0.664	0.668	-0.60%	0.666	0.668	-0.24%	0.666	0.667	-0.11%
	$[0.8, 0.4]$	W	0.245	0.254	-3.44%	0.193	0.193	-0.18%	0.177	0.177	-0.02%
		δ_1	0.760	0.737	3.05%	0.783	0.780	0.39%	0.795	0.793	0.29%
	$[0.5, 1]$	W	0.249	0.264	-6.19%	0.195	0.198	-1.57%	0.178	0.178	-0.17%
		δ_1	0.570	0.637	-11.89%	0.533	0.551	-3.39%	0.512	0.517	-1.10%
	$[0.5, 0.5]$	W	0.295	0.293	0.64%	0.248	0.249	-0.29%	0.233	0.232	0.44%
		δ_1	0.665	0.666	-0.16%	0.664	0.668	-0.65%	0.668	0.670	-0.27%
	$[0.4, 0.8]$	W	0.296	0.311	-5.11%	0.239	0.245	-2.57%	0.220	0.222	-0.94%
		δ_1	0.564	0.620	-10.05%	0.528	0.542	-2.60%	0.511	0.516	-1.09%
	$[0.4, 0.4]$	W	0.357	0.358	-0.25%	0.305	0.307	-0.64%	0.292	0.291	0.26%
		δ_1	0.666	0.665	0.02%	0.666	0.666	-0.03%	0.666	0.668	-0.29%

Table 1: Values and relative gaps between simulated and estimated waiting times and boarding probabilities under different scenarios.

4. Synthesis and conclusions

We have proposed a stochastic equilibrium model for transit assignment that includes capacity constraints (congestion at bus stops). The stochastic aspect of the model is incorporated through passenger decisions of boarding a specific bus of certain line at the bus stop. In the classical approach of common lines, parallel lines running between the same origin destination pair, passengers chose among this set of bus lines those that minimize expected travel time to their destination. In our stochastic approach, the decision of choosing a line becomes stochastic. For a bus service, the probability of being chosen depends on the difference between travel time in this line and overall expected travel time. Thus, every line has a chance to be selected, no matter how bad its service is. Following this approach we have defined Stochastic Common-line Equilibrium and we have proved its existence under mild conditions.

We have proposed as well a stochastic extension of Transit Equilibrium (Cominetti and Correa, 2001; Cepeda et al., 2006). The modelling approach assumes that passengers travel according to shortest hyperpaths to accomplish their origin-destination trips. Bus services are characterized by their observed frequency at the bus stops and their travel time to the next stop. The formulation is an extension of the hyperpath model for the deterministic case. Our definition of Stochastic Transit Equilibrium is then a straightforward adaptation of the definition of Transit Equilibrium. We characterize the equilibrium as a vector of feasible flows and expected travel times that must satisfy a set of simultaneous stochastic common-line problems coupled by flow conservation constraints.

When defining Stochastic Transit Equilibrium we have incorporated two new expressions for expected waiting time and network loads. In the definition of the common-line problems we introduce new generic network load distributions that now take into account not only effective frequency of the service but also the probability of boarding. The resulting expressions are hard to be obtained analytically by studying the stochastic process occurring at the bus stop; and are therefore supported, in the last part of the paper, through simulation of the embedded queueing process. In the simulations we have included both effects: overcrowding and stochastic behaviour at boarding. We found small differences between the simulation experiments and the expected results of the stochastic model (in terms of waiting time as well as transit lines loads), validating the correctness of the proposed formulation. It is important to note that the effect of overcrowding at bus stops (high congestion scenarios) is not significantly different from the deterministic approach which is widely accepted in the literature (see details in Table 1). Thus, incorporating the stochastic behavior of passengers does not affect the validity of our formulation of equilibrium.

An advantage of our formulation is that it is no longer necessary to solve dynamic programming problems to calculate expected travel times, which are now obtained through functional forms. A further objective is to incorporate this formulation to an integrated scheme of stochastic private and transit equilibrium (Baillon and Cominetti, 2008).

References

Baillon, J.-B., Cominetti, R., January 2008. Markovian traffic equilibrium. Mathematical Programming 111 (1-2), 33–56.

Bouzaiene-Ayari, B., Gendreau, M., Nguyen, S., 2001. Modeling bus stops in transit networks: A survey and new formulations. Transportation Science 35 (3), 304–321.

Cepeda, M., Cominetti, R., Florian, M., July 2006. A frequency-based assignment model for congested transit networks with strict capacity constraints: characterization and computation of equilibria. Transportation Research Part B: Methodological 40 (6), 437–459.

Chriqui, C., Robillard, P., 1975. Common bus lines. Transportation Scence 9, 115–121.

Cominetti, R., Correa, J., August 2001. Common-lines and passenger assignment in congested transit networks. Transportation Science 35 (3), 250–267.

De Cea, J., Fernández, E., 1993. Transit assignment for congested public transport systems: An equilibrium model. Transportation Science 27 (2), 133–147.

Gendreau, M., 1984. tude approfondie d'un modle d'quilibre pour l'affectation des passagers dans les rseaux de transport en commun. Ph.D. thesis, Departement dinformatique et de recherche operationnelle, Universit de Montral, Quebec.

Nguyen, S., Pallottino, S., November 1988. Equilibrium traffic assignment for large scale transit networks. European Journal of Operational Research 37 (2), 176–186.

Nguyen, S., Pallottino, S., 1989. Hyperpaths and shortest hyperpaths. In: Simeone, B. (Ed.), Lecture Notes in Mathematics. Vol. 1403. Springer-Verlag, pp. 258–271.

Nguyen, S., Pallottino, S., Gendreau, M., 1998. Implicit Enumeration of Hyperpaths in a Logit Model for Transit Networks. Transportation Science 32 (1), 54–64.

Schmcker, J.-D., Bell, M. G., Kurauchi, F., 2008. A quasi-dynamic capacity constrained frequency-based transit assignment model. Transportation Research Part B: Methodological 42 (10), 925 – 945.

Simon, C. P., Blume, L., 1994. Mathematics For Economists. W. W. Norton & Company, Inc., New York.

Spiess, H., 1984. Contributions a la theorie et aux outils de planification des reseaux de transport urbain. Ph.D. thesis, Departement dinformatique et de recherche operationnelle, Universit de Montral, Quebec.

Spiess, H., Florian, M., 1989. Optimal strategies: a new assignment model for transit networks. Transportation Research. Part B : Methodological 23 (2), 83–102.

Wu, J. H., Florian, M., Marcotte, P., 1994. Transit equilibrium assignment: A model and solution algorithms. Transportation Science 28 (3), 193–203.

Appendix A. Stationary distribution

In this appendix we develop the stationary distribution $\Pi = \{\pi_k\}_{k\geq 0}$ characterized by the balance equations obtained from the solution of the system $\Theta^T \Pi = \Pi\,(v+\mu)$ with $\theta_{0,0} = \mu$. The developments are based on the transition rates for the stochastic model summarized in section 3.

To synthesize the notation of the probability expressions in (10) and (11), hereafter, let us denote $\mathbb{P}(\text{board } j|\text{stay } k) \equiv \mathbb{P}(S = j|k)$ and $\mathbb{P}(\text{board} \geq j|\text{stay } k) \equiv \mathbb{P}(S \geq j|k)$

Notice that if $p = 1$ we recover the model by Cominetti and Correa (2001). Analytically

$$\theta_{k,0} = \mu \sum_{l=k}^{K} q_l, \quad \theta_{k,k-j} = \mu\,[0 + 1 \cdot q_j] = \mu q_j.$$

The following step is to compute the transition probabilities to formulate the balance equations. For that, we have to explicitly find the transition rates $\theta_{i,j}$. For this, let us define

$$\theta_{k,\bullet} = \sum_{j\neq k} \theta_{k,j}.$$

At this stage, what we need to calculate is $\lambda = \theta_{k,\bullet} + \theta_{k,k}$.

The deterministic assignment can be modelled assuming $p = 1$; in such a case, we can write the following expressions. For $k = 0$,

$$\theta_{0,\bullet} = v, \quad \text{therefore} \quad \theta_{0,0} = -v \quad \text{and then} \quad \rho_{0,0} = 1 + \frac{1}{\lambda}(-v);$$

and for $k \geq 1$

$$\theta_{k,\bullet} + \theta_{k,k} = \underbrace{\mu \sum_{l=k}^{K} q_l}_{\theta_{k,0}} + \underbrace{\mu \sum_{j=0}^{k-1} q_j}_{\sum_{j=0}^{k-1} \theta_{k,k-j}} + \underbrace{v}_{\theta_{k,k+1}} = \mu + v.$$

Therefore $\lambda = \mu + v$, and then

$$\begin{aligned} \rho_{0,0} &= 1 + \frac{1}{\mu+v}(-v) = \frac{v+\mu-v}{\mu+v} \qquad = \frac{\mu}{\mu+v}; \\ \rho_{i,j} &= \frac{\theta_{i,j}}{v+\mu}, \qquad\qquad \text{if } (i,j) \neq (0,0). \end{aligned}$$

Let us now generalize the deterministic assignment model to include the stochastic behaviour in passenger decisions, through $p < 1$. In this case, for $k = 0$,

$$\theta_{0,\bullet} = v, \quad \text{therefore} \quad \theta_{0,0} = -v \quad \text{and then} \quad \rho_{0,0} = 1 + \frac{1}{\lambda}(-v).$$

On the other hand, for $k \neq 0$ the transition rates can be computed as follows

$$\theta_{k,\bullet} + \theta_{k,k} = \mu p^k \sum_{l=k}^{K} q_l + \mu \sum_{j=0}^{k-1} \left[\sum_{l=j+1}^{K} q_l \mathbb{P}(S = j|k) + q_j \mathbb{P}(S \geq j|k) \right] + v. \tag{A.1}$$

Expression (A.1) synthesizes the generic case for k positive. Taking into account that this model incorporates capacity constraints on bus sizes, we have two options to compute such transition rates. Either $k > K$ or $1 \leq k \leq K$ ($j \leq k - 1 \leq K - 1 \leq K \rightsquigarrow j + 1 \leq K$), where as stated before, K denotes the physical capacity of a bus. Then, considering first that $k > K$, we have

$$\begin{aligned}
\theta_{k,\bullet} + \theta_{k,k} &= \mu \sum_{j=0}^{k-1} \left[\sum_{l=j+1}^{K} q_l \mathbb{P}(S = j|k) + q_j \mathbb{P}(S \geq j|k) \right] + v \\
&= \mu \underbrace{\sum_{j=0}^{k-1} \sum_{l=j+1}^{K} q_l \mathbb{P}(S = j|k)}_{= 0 \text{ si } j+1 > K} + \mu \sum_{j=0}^{k-1} \underbrace{q_j}_{= 0 \text{ si } j > K} \mathbb{P}(S \geq j|k) + v \\
&= \mu \sum_{j=0}^{K-1} \sum_{l=j+1}^{K} q_l \mathbb{P}(S = j|k) + \mu \sum_{j=0}^{K} q_j \mathbb{P}(S \geq j|k) + v \\
&= \mu \sum_{l=1}^{K} \sum_{j=0}^{l-1} q_l \mathbb{P}(S = j|k) + \mu \sum_{j=0}^{K} q_j \mathbb{P}(S \geq j|k) + v \\
&= \mu \sum_{l=1}^{K} q_l \left(\sum_{j=0}^{l-1} \mathbb{P}(S = j|k) + \mathbb{P}(S \geq l|k) \right) + \mu q_0 \underbrace{\mathbb{P}(S \geq 0|k)}_{=1} + v \\
&= \mu \sum_{l=1}^{K} q_l \underbrace{(\mathbb{P}(S < l|k) + \mathbb{P}(S \geq l|k))}_{=1} + \mu q_0 + v \\
&= v + \mu (1 - q_0 + q_0) \\
&= v + \mu.
\end{aligned}$$

The other option for computing the transition rates in expression (A.1) is the case in which $1 \leq k \leq K$ ($j \leq k-1 \leq K-1 \leq K \rightsquigarrow j+1 \leq K$). Then,

$$\begin{aligned}
\theta_{k,\bullet} + \theta_{k,k} &= v + \mu p^k \sum_{l=k}^{K} q_l + \mu \sum_{j=0}^{k-1} \sum_{l=j+1}^{K} q_l \mathbb{P}(S=j|k) + \mu \sum_{j=0}^{k-1} q_j \mathbb{P}(S \geq j|k) \\
&= v + \mu p^k \sum_{l=k}^{K} q_l + \mu \sum_{l=1}^{K} \sum_{j=0}^{\min\{l-1,k-1\}} q_l \mathbb{P}(S=j|k) + \mu \sum_{j=0}^{k-1} q_j \mathbb{P}(S \geq j|k) \\
&= v + \mu p^k \sum_{l=k}^{K} q_l + \mu \sum_{l=1}^{k-1} \sum_{j=0}^{l-1} q_l \mathbb{P}(S=j|k) + \mu \sum_{l=k}^{K} \sum_{j=0}^{k-1} q_l \mathbb{P}(S=j|k) + \mu \sum_{l=0}^{k-1} q_l \mathbb{P}(S \geq l|k) \\
&= v + \mu p^k \sum_{l=k}^{K} q_l + \mu \sum_{l=1}^{k-1} q_l \mathbb{P}(S<l|k) + \mu \sum_{l=k}^{K} q_l \mathbb{P}(S<k|k) + \mu \sum_{l=0}^{k-1} q_l \mathbb{P}(S \geq l|k) \\
&= v + \mu \sum_{l=k}^{K} q_l \underbrace{\left(p^k + \mathbb{P}(S<k|k)\right)}_{=1} + \mu \sum_{l=1}^{k-1} q_l \underbrace{\left(\mathbb{P}(S<l|k) + \mathbb{P}(S \geq l|k)\right)}_{=1} + \mu q_0 \underbrace{\mathbb{P}(S \geq 0|k)}_{=1} \\
&= v + \mu \sum_{l=0}^{K} q_l \\
&= v + \mu.
\end{aligned}$$

Therefore, we can say that independently of the capacity constraint given indirectly by K, when $k > 0$, $\lambda = \mu + v$. Hence, the transition probabilities are calculated as follows

$$\begin{aligned}
\rho_{0,0} &= 1 - \frac{v}{\mu+v} = \frac{v+\mu-v}{\mu+v} \quad = \frac{\mu}{\mu+v}; \\
\rho_{i,j} &= \frac{\theta_{i,j}}{v+\mu}, \qquad \text{if } (i,j) \neq (0,0)
\end{aligned}$$

Then, the solution of the system $P^T\Pi = \Pi \rightsquigarrow \Theta^T\Pi = \Pi\,(v+\mu)$ considering $\theta_{0,0} = \mu$, is as follows. For $k=0$ we have:

$$\begin{aligned}
(v+\mu)\,\pi_0 &= \sum_{k=0}^{\infty} \pi_k \theta_{k,0} = \mu\pi_0 + \sum_{k=1}^{\infty} \pi_k \mu p^k \sum_{l=k}^{K} q_l; \\
v\pi_0 &= \mu \sum_{k=1}^{\infty} \pi_k p^k \sum_{l=k}^{K} q_l = \mu \sum_{l=k}^{K} q_l \sum_{k=1}^{l} \pi_k p^k; \\
v\pi_0 &= \mu \sum_{l=1}^{K} q_l \sum_{k=1}^{l} \pi_k p^k; \qquad \text{(A.2)}
\end{aligned}$$

and for $k > 0$

$$
\begin{aligned}
(v+\mu)\,\pi_k &= \sum_{k'=0}^{\infty} \pi_{k'}\theta_{k',k} = \pi_{k-1}v + \sum_{k'=k}^{\infty} \pi_{k'}\theta_{k',k} = \pi_{k-1}v + \sum_{j=0}^{\infty} \pi_{k+j}\theta_{k+j,k} \\
&= \pi_{k-1}v + \sum_{j=0}^{K} \pi_{k+j}\theta_{k+j,k}. \qquad \text{(A.3)}
\end{aligned}
$$

By manipulating terms and making explicit the θ expressions, we get the following development

$$
\begin{aligned}
&= \pi_{k-1}v + \sum_{j=0}^{K} \pi_{k+j}\mu\left[\mathbb{P}(C > j)\,\mathbb{P}(S = j|k+j) + \mathbb{P}(C = j)\,\mathbb{P}(S \geq j|k+j)\right] \\
&= \pi_{k-1}v + \mu\sum_{j=0}^{K} \pi_{k+j}\left[\sum_{l=j+1}^{K} q_l\mathbb{P}(S = j|k+j) + q_j\mathbb{P}(S \geq j|k+j)\right] \\
&= \pi_{k-1}v + \mu\left[\sum_{l=1}^{K} q_l \sum_{j=0}^{l-1} \pi_{k+j}\mathbb{P}(S = j|k+j) + \sum_{j=0}^{K} q_j\pi_{k+j}\mathbb{P}(S \geq j|k+j)\right] \\
&= \pi_{k-1}v + \mu\sum_{l=1}^{K} q_l\left[\sum_{j=0}^{l-1} \pi_{k+j}\mathbb{P}(S = j|k+j) + \pi_{k+l}\mathbb{P}(S \geq l|k+l)\right] + q_0\pi_k \\
&= \pi_{k-1}v + \mu\sum_{l=1}^{K} q_l\left[\sum_{j=0}^{l} \pi_{k+j}\mathbb{P}(S = j|k+j) + \pi_{k+l}\mathbb{P}(S > l|k+l)\right] + q_0\pi_k.
\end{aligned}
$$

Thus, expression (A.3) becomes

$$
(v+\mu)\,\pi_k = \pi_{k-1}v + \mu\sum_{l=0}^{K} q_l\left[\sum_{j=0}^{l} \pi_{k+j}\mathbb{P}(S = j|k+j) + \pi_{k+l}\mathbb{P}(S > l|k+l)\right]. \qquad \text{(A.4)}
$$

Finally, equations (A.2) and (A.4) define the following system

$$
\begin{aligned}
v\pi_0 &= \mu\sum_{l=1}^{K} q_l \sum_{k=1}^{l} \pi_k p^k; \\
(v+\mu)\,\pi_k &= \pi_{k-1}v + \mu\sum_{l=0}^{K} q_l\left[\sum_{j=0}^{l} \pi_{k+j}\mathbb{P}(S = j|k+j) + \pi_{k+l}\mathbb{P}(S > l|k+l)\right], \qquad \forall k \geq 1.
\end{aligned}
$$