Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 00 (2017) 000–000

Transportation Research Procedia

www.elsevier.com/locate/procedia

INTERNATIONAL SYMPOSIUM ON TRAFFIC AND TRANSPORTATION THEORY

ISTTT22

Stochastic dynamic switching in fixed and flexible transit services as market entry-exit real options

Qian-Wen Guo[a], Joseph Y. J. Chow[b,*], Paul Schonfeld[c]

[a]Department of Finance and Investment, Sun Yat-Sen University, Bldg 388, 135 Xingangxi Rd, Guangzhou 510275, China
[b]Department of Civil and Urban Engineering, New York University, 6 Metro Tech Center, Brooklyn, NY 11201, USA
[c]Department of Civil and Environmental Engineering, University of Maryland, 1173 Glenn Martin Hall, College Park, MD 20742, USA

Abstract

The first analytical stochastic and dynamic model for optimizing transit service switching is proposed for "smart transit" applications and for operating shared autonomous transit fleets. The model assumes a region that requires many-to-one last mile transit service either with fixed-route buses or flexible-route, on-demand buses. The demand density evolves continuously over time as an Ornstein-Uhlenbeck process. The optimal policy is determined by solving the switching problem as a market entry and exit real options model. Analysis using the model on a benchmark computational example illustrates the presence of a hysteresis effect, an indifference band that is sensitive to transportation system state and demand parameters, as well as the presence of switching thresholds that exhibit asymmetric sensitivities to transportation system conditions. The proposed policy is computationally compared in a 24-hour simulation to a "perfect information" set of decisions and a myopic policy that has been dominant in the flexible transit literature, with results that suggest the proposed policy can reduce by up to 72% of the excess cost in the myopic policy. Computational experiments of the "modular vehicle" policy demonstrate the existence of an option premium for having flexibility to switch between two vehicle sizes.

Peer-review under responsibility of the scientific committee of the 22nd International Symposium on Transportation and Traffic Theory.

Keywords: public transit; market entry and exit real options; stochastic dynamic optimization; flexible transit; last mile problem

* Corresponding author. Tel.: +1-646-997-4030; fax: +1-646-997-3433.
E-mail address: joseph.chow@nyu.edu

Peer-review under responsibility of the scientific committee of the 22nd International Symposium on Transportation and Traffic Theory.

Nomenclature

(subscript c denotes conventional/fixed transit; f denotes flexible transit)
L = length of region (miles)
W = width of region (miles)
J = line haul distance from terminal to the nearest corner of the region (miles)
N_c, N_f = number of routes/zones in region
F_c, F_f = fleet size in region
$S_c(t), S_f(t)$ = vehicle size in region (seats/vehicle)
$h_c(t), h_f(t)$ = headway in region (hours)
V_c, V_f = bus operating speed within the region (miles/hour)
V_x = average passenger access speed (miles/hour)
$C_{sc}(t), C_{sf}(t)$ = total cost accrued from time t to $t + dt$ in region ($)
a = fixed cost of bus operation per unit time ($/hour)
b = variable cost of bus operation per unit time and seat ($/seat·hour)
v_i, v_w, v_x = users' value of time for in-vehicle time (i), wait time (w), and access time (x) ($/passenger hour)
y_c, y_f = ratio of direct line haul non-stop speed over local speed
z_c, z_f = ratio of local non-stop speed over local speed
c_c, c_f = average total cost per trip ($/trip)
ρ = discount rate

Demand parameters
$Q(t)$ = demand density at time t (trips/sq.mile·hour)
μ = mean reversion coefficient for demand process
m = stationary demand density (trips/sq.mile·hour)
σ = process volatility for demand process

Fixed/conventional transit only
r = route spacing for region (miles)
d = bus stop spacing (miles)
D_c = round trip bus distance (miles)

Flexible transit only
D_0 = distance of one flexible bus tour within a service zone (miles)
D_f = equivalent line haul distance for flexible bus (miles)
n = number of passengers in one flexible bus tour
k = constant for a grid network (Daganzo, 1984a) for flexible bus
A = service zone area for flexible bus (sq. miles)
u = average number of passengers per stop for flexible bus

1. Introduction

The potential of optimal timing and control of transit systems under uncertainty continues to grow in today's data-driven environment. There are countless examples of such applications, including: determining when to allow fixed-route services to deviate; optimal holding strategies for buses; adjusting size of vehicle groups (e.g. trains) that are dispatched; positioning idle on-demand vehicles; and determining "price surges". However, there are very few fundamentally general analytical methods available to time decisions under dynamic uncertainty in this domain. In this study, we explore one such timing method based on real options theory, and evaluate its effectiveness in well-studied problems of time-dependent changes between two different transit fleet operating modes.

It has long been known that different demand density levels warrant transit services with different operating policies and degrees of flexibility (Saltzman, 1973; Jacobson, 1980; Adebisi and Hurdle, 1982). Some studies sought to determine thresholds based on demand densities between different transit services, including fixed-route and flexible-route systems (Daganzo, 1984b; Chang and Schonfeld, 1991a; Quadrifoglio and Li, 2009; Qiu et al, 2015). Several studies have examined the problem of integrating fixed-route and flexible-route transit, primarily under the many-to-one service setting which applies either to last mile service design or to monocentric city structures (Chang and Schonfeld, 1991b; Kim and Schonfeld, 2013, 2014). These include joint design of fixed transit lines and feeder services (e.g. Kim and Schonfeld, 2014). With advances in artificial intelligence and autonomous vehicle technologies, as illustrated in Fig. 1, and with planned deployments in Dubai (Spera, 2016) and Singapore

(Ackerman, 2016), algorithms for optimal control of fleet systems over time are more urgent than ever. For example, autonomous vehicle fleets may be dynamically switched between fixed-route and on-demand operations.

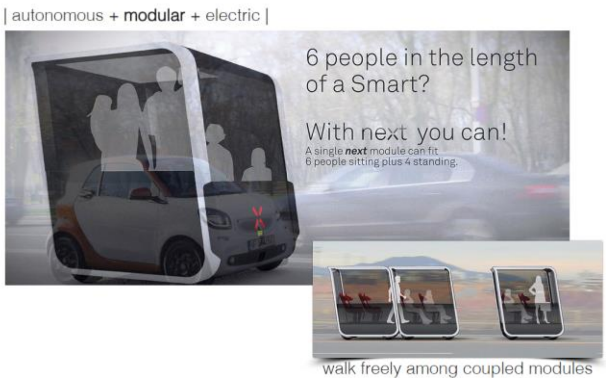

Fig. 1. Illustration of shared autonomous fleets with modular vehicle size (source: www.next-future-mobility.com).

The literature includes several studies for optimizing the assignment of transit vehicles to fixed or flexible services. Kim and Schonfeld (2012) proposed an analytical model framework for comparing operations among fixed-route only, flexible-route only, and an integrated service that temporally switches between the two during peak and off-peak periods. Quadrifoglio and Li (2009) and Qiu et al. (2015) examined analytical models of transit services that deviate from fixed routes to provide flexible drop-offs, and where to deterministically switch between them (Li and Quadrifoglio, 2010). Kim and Schonfeld (2013 and 2014) proposed an integrated service model that can alternate service types over both time and multiple sub-regions. Some studies have also explored with optimization models how mixed fleets consisting of multiple vehicle types or sizes should be allocated among various transit services (Lee et al, 1995; Fu and Ishkanov, 2004), and how such mixed fleets should be switched between different transportation services at different times (Kim and Schonfeld, 2013).

All these integrated service options are static policies since they are not designed to adapt to new information, and thus fail to exploit advances in increasingly pervasive information and communications technologies (ICTs). Dynamic flexible transit services (Djavadian and Chow, 2016) have become such a viable alternative for serving passengers that many new service providers have cropped up in the private sector alone: e.g. Uber, Lyft, Via, RideCo, and Bridj. In this context, there have been studies optimizing the dynamic routing (see Psaraftis et al., 2016), dynamic pricing (Sayarshad and Chow, 2015), dynamic vehicle waiting strategies (Thomas, 2007), and dynamic relocation of idle taxis (e.g. Yuan et al., 2011). Studies have not looked at the problem of dynamically allocating or switching vehicles between fixed and flexible routes under time-variant uncertainty.

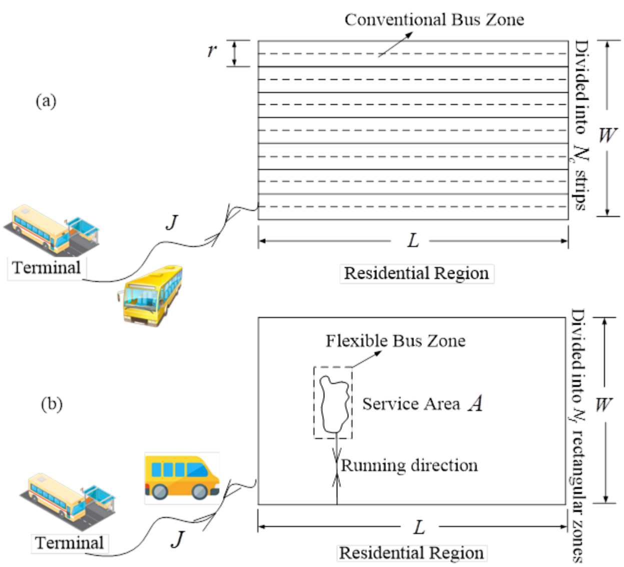

Fig. 2. (a) Fixed-route and (b) flexible route last mile transit service.

We propose to modify the static fixed/flexible service policy in the literature into a stochastic, dynamic policy in a many-to-one (M-to-1) system. This very adaptable M-to-1 structure, shown in Fig. 2, supports many applications: monocentric city designs, last mile problem, planned event logistics, and multimodal infrastructure planning, to name a few. As noted in Chang and Schonfeld (1991a), multiple such structures that are connected at a central terminal can serve many-to-many (M-to-M) demand patterns, thus connecting all possible origin destination pairs in an urban region with at most one transfer. However this study focuses on the M-to-1 system. For example, this setting includes having a fixed-route trunk transit service ending at the Terminal in Fig. 2, followed by either fixed-route buses, as shown in Fig. 2(a), or some flexible bus or shuttle service like UberPOOL in Fig. 2(b).

There are numerous methods for finding optimal dynamic policies, divided between analytical methods and numerical methods. Analytical methods with discrete stochastic variables include dynamic programming using the Bellman Equation (see Powell, 2011). Solution methods for continuous stochastic processes can be found in the real options literature (see Dixit and Pindyck, 1994). For example, Li et al. (2015) studied the transit technology timing problem using an analytical real options model.

In cases where the policy is too complex to optimize analytically, either due to the curse of dimensionality or to additional dependencies between variables, numerical methods have been employed. Many such methods are called approximate dynamic programming (Powell, 2011), and some involve the use of least squares Monte Carlo simulation (Chow and Regan, 2011a,b; Chow et al., 2011; Chow and Sayarshad, 2015).

Because the underlying cost valuation model for fixed and flexible service policies is analytical, we propose using an analytical method to find the optimal policy as well. In the model, ridership demand is a continuous stochastic variable. As such, analytical real options methods will be used to quantify options for switching between modes. The option to switch between modes has been called a market entry and exit model by Dixit (1989) for geometric Brownian motion processes, and further extended to mean-reverting processes by Sødal et al. (2008), Metcalf and Hassett (1995), Sarkar (2003), and Tsekrekos and Yannacopoulos (2016)). This approach has many applications in transportation and shipping (Sødal et. al., 2008, Sødal et. al., 2009, Gkochari, 2015) as well as in other fields, such as manufacturing and logistics (Cadenillas et. al., 2010), electricity networks (Parpa and Webster, 2014), and natural resource management (Brennan & Schwartz, 1985).

This study is the first to formulate a stochastic dynamic policy for switching service systematically between fixed and flexible services with an analytical model. It can be used to coordinate partnerships between private operators and transit agencies, and to manage autonomous fleets such as the Next Future Mobility system shown in Fig.1. The remainder of the study is divided into a Section 2 on defining the problem and reviewing the analytical methods to be used, Section 3 on the proposed model, Section 4 on the model properties using a benchmark example, Section 5 on numerical validation of the method and its application to modular shared autonomous vehicle fleets, and concluding in Section 6.

2. Problem definition

2.1. Problem illustration

An illustration is presented to motivate the formal definition. Consider a system with random demand Q_t (discrete version of $Q(t)$) over multiple discrete periods t, where the demand fluctuates between 0 and 10. There are two modes of operation: a flexible transit mode with a linear cost function $C_{sc} = 14 + 0.5Q_t$, and a fixed-route transit mode with a cost function $C_{sf} = 8 + 1.5Q_t$. The critical density where one service outperforms the other is 6.00, where fixed service is preferred if density exceeds the threshold. One simulated outcome of this system operating over 20 discrete periods is shown in Table 1 and Fig. 3, along with the threshold for switching between fixed and flexible transit. If the demand is absolutely deterministic or known beforehand, and there is no cost to switching from one service to the other, then the operator should run a fixed service whenever demand is above the dashed line, and a flexible service otherwise. There are nine periods that warrant fixed service: 1, 2, 3, 4, 8, 10, 13, 14, and 18. If the initial service at time zero is a fixed service, switching would occur nine times.

Table 1. Illustration parameters.

t	1	2	3	4	5	6	7	8	9	10
Q_t	8.77	7.50	6.37	6.50	1.33	1.51	1.48	6.53	0.44	9.78
C_{sc}	18.39	17.75	17.19	17.25	14.67	14.76	14.74	17.27	14.22	18.89
C_{sf}	21.16	19.25	17.56	17.75	10.00	10.27	10.22	17.80	8.66	22.67
t	11	12	13	14	15	16	17	18	19	20
Q_t	4.34	0.92	8.09	7.42	5.61	3.11	1.14	6.34	0.70	3.43
C_{sc}	16.17	14.46	18.05	17.71	16.81	15.56	14.57	17.17	14.35	15.72
C_{sf}	14.51	9.38	20.14	19.13	16.42	12.67	9.71	17.51	9.05	13.15

In this deterministic setting, the total cost under the optimal switching without switching cost is simply the sum of the minima of the two service costs over all periods, leading to total cost of 283.64. If there is a symmetric (i.e. equal in either direction) switching cost of 0.5, the optimal decision is to withhold switching in periods 8 and 18 as the cost of switching to and from the fixed service would exceed the operational cost savings. That leaves 5 switches in place. The resulting operations would have operating costs of 338.07 plus switching costs of 2.5 (from 5 switches) for a total of 340.57. If the symmetric switching cost increases further to 1.8, the optimal decision is to forsake switching to fixed service in periods 13 and 14. Note that decisions in the deterministic setting involve looking at trade-offs over multiple periods. For example, period 13 actually shows a cost saving of 2.09 for fixed transit, which exceeds the switching cost of 1.8, but because period 14 only has a cost saving of 1.42 before switching back to flexible transit, the cumulative cost savings are less than the switching cost to “enter” and “exit” the fixed transit mode. This problem can be solved as a dynamic programming problem in the deterministic case.

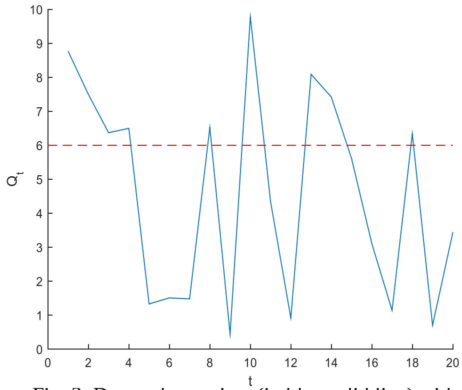

Fig. 3. Demand over time (in blue solid line) with switching threshold (orange dashed line).

However, in deciding whether to switch service in the stochastic dynamic context, one may only have historical information about the demand up to some period t (which may also be continuous) and information on the current state of the system (whether it is currently providing fixed or flexible service. The decision is not just dependent on the historical data, but also on the expectation of the state of the system in the near future. In the example, if the operator was trying to dynamically determine whether to switch from flexible to fixed service at period 13, the switching cost at that time is below the operational cost savings. However, an expectation of where the demand will go may show that the expected savings minus the cost of entry and exit would be negative. This is the problem that we formulate and solve.

2.2. Formal definition

An urban system is assumed to be geographically represented with a single rectangular region of length L and width W, and connected to a hub via line haul of length J, as shown in Fig. 2, which is adapted from Kim and Schonfeld (2014). The demand density is assumed to be uniformly distributed over the region. If there are great variations in the region, it can be further split into multiple regions where each has fairly homogeneous demand. The demand is assumed to be predominantly M-to-1 from each region to the hub (and 1-to-M in the opposite direction). This design accommodates a wide range of configurations, such as a hub located in the middle of a rectangular region serving four quadrants as a monocentric city, or breaking down one rectangular region into multiple sub-regions that are each served by either fixed or flexible service.

In the fixed-route conventional mode, the region is subdivided into N_c routes of width r and length L. Buses operating in this mode would travel from the terminal across the line haul length J to the corner, up along the edge of the region to the designated strip, and then make a round trip along the length of L to pick up and drop off passengers. In the flexible service mode, the regions are subdivided into a grid of N_f zones of area A.

In the formal problem, we adopt a continuous time assumption as opposed to the discrete periods from the illustration earlier. The two services are assumed to behave consistently with the assumptions from Chang and Schonfeld (1991a); there exists a critical demand density such that fixed transit operates at a lower cost for higher demand and flexible transit for lower demand.

The demand density $Q(t)$ is assumed to be a time-dependent stochastic process, where real-time information is obtained from smart cards, mobile ticketing, or traffic information systems. A mean-reverting Ornstein-Uhlenbeck (O-U) process (Sarkar, 2003) is assumed, as expressed in Eq. (1):

$$dQ = \mu(m - Q)dt + \sigma Q dw \qquad (1)$$

where μ is a mean reversion coefficient, m is the long-term demand density, σ is the process volatility, and $dw \sim N(0, dt)$ is an increment in the Wiener process. The solution to Eq. (1) is given by Eq. (2).

$$Q(t) = Q(0)\exp\left(-\left(\mu + \frac{1}{2}\sigma^2\right)t + \sigma w\right) + \frac{\mu m}{\mu + \frac{1}{2}\sigma^2}\left(1 - \exp\left(-\left(\mu + \frac{1}{2}\sigma^2\right)t\right)\right) \tag{2}$$

The expected value and conditional variance of $Q(t)$ are shown in Eq. (3) and (4) (Tsekrekos, 2010), respectively.

$$E_0[Q(t)] = m + (Q(0) - m)\exp(-\mu t) \tag{3}$$

$$\begin{aligned} Var_0[Q(t)] = {} & \frac{m^2\sigma^2}{2\mu - \sigma^2} - \exp(-2\mu t)\,(Q(0) - m)^2 + \frac{2m\sigma^2(Q(0) - m)}{\mu - \sigma^2}\exp(-\mu t) \\ & + \frac{2\mu^2(Q(0) - m)^2 - \mu\sigma^2 Q(0)(3Q(0) - 2m) + \sigma^4 Q^2(0)}{(\mu - \sigma^2)(2\mu - \sigma^2)}\exp(-(2\mu - \sigma^2)t) \end{aligned} \tag{4}$$

O-U processes are generalizations of the geometric Brownian motion (GBM) that allow for stationarity, and have been shown to fit the data in investment models better than GBM despite the increased complexity (Tsekrekos, 2010).

Several estimation methods are available for estimating parameters of the Ornstein-Uhlenbeck process, such as a maximum likelihood approach (Zagoraiou and Antognini 2009), an application of the Fortet integral equation (Ditlevsen and Ditlevsen 2008), and least square estimation (Hu and Nualart 2010). An example maximum likelihood estimation of O-U parameters is conducted by Chow and Regan (2011c).

The formal problem is set at some time t, when the operator observes $Q(s)$ for $s < t$ and a current operating mode function $a(t) \in \{0,1\}$ where a value of 0 indicates flexible service and 1 indicates fixed service. The operator decides whether to switch from $a(t - dt) = 0$ (or 1) in the previous time increment dt to $a(t) = 1$ (or 0), where the operating mode impacts the incremental operating cost $\Phi(a(t))$ and future expected value of the system. The problem is to determine an optimal policy such that total realized operating cost plus switching cost over an entire daily cycle is expected to be minimized. Decisions made on service type are assumed to occur instantaneously (or negligibly so with respect to the time horizon), and switching from one service to another would incur a switching. The optimal policy problem can be expressed in terms of a Bellman Equation in Eq. (2), where the demand $Q(t)$ is from Eq. (1), V is the policy value, and ρ is a discount factor over an increment in time.

$$V(Q(t)) = \max_{\mathrm{a(t)}}\left(\Phi(a(t)) + \rho E\left[V(Q(t + dt))\middle|Q(t), a(t)\right]\right) \tag{5}$$

3. Proposed model

To solve for the optimal policy, we model the incremental operational cost functions using analytical expressions for fixed and flexible transit services as functions of demand. The functions are used to evaluate current and future expectations of the cost based on the current state of the demand process and operating mode. Since there is a single stochastic process and two services based on it, the problem falls under the category of market entry-exit models under O-U process, and is solved as such.

3.1. Incremental cost function

The first step is to construct an expression for the incremental cost function $\Phi(a(t))$. The following expressions are based on Chang and Schonfeld (1991a), and readers are referred there for derivation details.

3.1.1. Fixed transit service cost

The conventional bus cost C_{sc} is the sum of the bus operating cost C_o, user in-vehicle cost C_v, user waiting cost C_w, and user access cost C_x.

$$C_{sc} = C_o + C_v + C_w + C_x \tag{6}$$

Eq. (6) has been shown by Kim and Schonfeld (2013) to expand to the following Eq. (7).

$$C_{sc} = \frac{LWQD_c(a+bS_c)}{V_cS_c} + \frac{v_iLWMQ}{V_c} + \frac{v_wS_cW}{2r} + \frac{v_xLWQ}{4V_x}(r+d) \tag{7}$$

The fleet size F_c is defined from $F_c = \frac{WD_c}{V_crh_c}$, the headway is $h_c = \frac{S_c}{LrQ}$, the average bus round trip distance is $D_c = \frac{2J}{y_c} + \frac{W}{z_c} + 2L$, and the average user trip distance is $M = \frac{J}{y_c} + \frac{W}{2z_c} + \frac{L}{2}$. The average cost per trip c_c then equals total cost divided by total trip demand LWQ, as shown in Eq. (8).

$$c_c = \frac{D_c(a+bS_c)}{V_cS_c} + \frac{v_iM}{V_c} + \frac{v_wS_c}{2rLQ} + \frac{v_x}{4V_x}(r+d) \tag{8}$$

Using first-order conditions of the average cost per trip c_c with regard to route spacing r and vehicle size S_c, we obtain the optimal route spacing and vehicle size in Eq. (9) and (10).

$$r^* = \left(\frac{8av_wV_x^2D_c}{v_x^2LQV_c}\right)^{\frac{1}{3}} \tag{9}$$

$$S_c^* = \left(\frac{8a^2V_xLQD_c^2}{v_xv_wV_c^2}\right)^{\frac{1}{3}} \tag{10}$$

Since the closed form solutions of route spacing r and vehicle size S_c are available, they can be inserted into Eq. (7), which becomes a function of the demand Q only in Eq. (11a).

$$C_{sc}^*(Q) = 3LWQ\left(\frac{v_wv_xaD_c}{8V_xV_cLQ}\right)^{\frac{1}{3}} + \frac{bLWQD_c}{V_c} + \frac{v_xLWQd}{4V_x} + \frac{v_iLWQM}{V_c} \tag{11a}$$

If S_c is fixed, we get Eq. (11b).

$$C_{sc}^*(Q;S_c) = \frac{LWQD_c(a+bS_c)}{V_cS_c} + \frac{v_iLWMQ}{V_c} + W\sqrt{\frac{v_wS_cv_xLQ}{2V_x}} + + \frac{v_xLWQ}{4V_x}\left(\sqrt{\frac{2V_xv_wS_c}{v_xLQ}} + d\right) \tag{11b}$$

3.1.2. Flexible transit service cost

In the flexible service mode, the region is further divided into a grid of N_f identical size zones with area A. A vehicle travels directly from the hub to a zone and then performs a tour to collect all passengers assigned to it before returning to the hub along the same path. Unlike for fixed-route service, the fleet size for flexible bus service, as shown in Eq. (12), depends on tour distance within a service zone, which in turn depends on demand, as shown in Eq. (13) per Stein (1978). The headway is $h_f = \frac{S_f}{AQ}$, which depends on vehicle size S_f. The equivalent round trip distance from the terminal to the service zone is $D_f = \frac{L+W}{z_f} + \frac{2J}{y_f}$.

$$F_f = \frac{D_0 + D_f}{V_fh_f} \tag{12}$$

$$D_0 = kA\sqrt{\frac{Qh_f}{u}} = \frac{kLW}{N_f}\sqrt{\frac{Qh_f}{u}} \tag{13}$$

The total cost is shown in Eq. (14), which includes the same cost components present in Eq. (6).

$$C_{sf} = \frac{LWD_fQ\left(a+bS_f\right)}{V_fS_f} + \frac{LWQk\sqrt{\frac{A}{uS_f}}\left(a+bS_f\right)}{V_f} + \frac{v_iLWD_fQ}{2V_f} + \frac{v_iLWQ\sqrt{\frac{AS_f}{u}}}{2V_f} + \frac{v_wS_fLW}{2A} \tag{14}$$

The average cost per trip c_f is provided in Eq. (15), and the optimal vehicle size and service zone area from first order conditions are provided in Eq. (16) and (17), respectively:

$$c_f = \frac{D_f\left(a+bS_f\right)}{V_fS_f} + \frac{k\sqrt{\frac{A}{uS_f}}\left(a+bS_f\right)}{V_f} + \frac{v_iD_f}{2V_f} + \frac{v_i\sqrt{\frac{AS_f}{u}}}{2V_f} + \frac{v_wS_f}{2AQ} \tag{15}$$

$$S_f^* = \left(\frac{ua^3D_f^3Q}{v_wk^2V_f\left(b+\frac{v_i}{2}\right)^2}\right)^{\frac{1}{5}} \tag{16}$$

$$A^* = \left(\frac{V_f^3v_w^3u^{\frac{8}{3}}aD_f^3}{Q^{\frac{7}{3}}k^4Y^{\frac{10}{3}}\left(b+\frac{v_i}{2}\right)^2}\right)^{\frac{1}{5}} \tag{17}$$

where Y is the intermediate variable represented in Eq. (18).

$$Y = \left(v_wa^2k^3V_f\right)^{\frac{1}{5}} + \left(uQD_f^3\left(b+\frac{v_i}{2}\right)^3\right)^{\frac{1}{5}} \tag{18}$$

The optimal zone area A^* and vehicle size S_f^* are substituted into Eq. (14) to obtain the total optimized cost of flexible bus service in Eq. (19a), which is a function of demand density Q.

$$C_{sf}^*(Q) = LWQ\left[\left(\frac{v_wa^2k^2D_f^2\left(b+\frac{v_i}{2}\right)^2}{V_f^2uQ}\right)^{\frac{1}{5}} + 1.5\left(\frac{v_wa^2k^2}{V_f^4}\right)^{\frac{1}{5}}\left(\frac{Y^2}{uQ}\right)^{\frac{1}{3}} + \frac{D_f\left(b+\frac{v_i}{2}\right)}{V_f}\right] \tag{19a}$$

If S_f is fixed, we obtain Eq. (19b).

$$\begin{aligned} C_{sf}^*\left(Q;S_f\right) = {} & \frac{LWD_fQ\left(a+bS_f\right)}{V_fS_f} + \frac{v_iLWD_fQ}{2V_f} + \frac{v_wS_fLW}{2}\left(\frac{v_wS_fV_f}{Qk\left(a+bS_f\right)}\sqrt{uS_f} + \frac{v_i}{2}\sqrt{\frac{S_f}{u}}\right)^{-\frac{2}{3}} \\ & + \frac{LWQ}{V_f}\left(\frac{k}{\sqrt{uS_f}}\left(a+bS_f\right) + \frac{v_i}{2}\sqrt{\frac{S_f}{u}}\right)\left(\frac{v_wS_fV_f}{Q\left(a+bS_f\right)k}\sqrt{uS_f} + \frac{v_i}{2}\sqrt{\frac{S_f}{u}}\right)^{\frac{1}{3}} \end{aligned} \tag{19b}$$

3.1.3. Incremental cost savings function

Based on the two cost functions, we can construct a cost function of the immediate cost savings accrued from time t to $t + dt$ when operating in one mode relative to the other. This is shown in Eq. (20).

$$\Phi\left(Q(t)\right) = C_{sf}^*\left(Q(t);S_f\right) - C_{sc}^*\left(Q(t);S_c\right) \tag{20}$$

3.2. Policy valuation

The following additional notation is introduced in this section pertaining to the optimal policy.

Additional policy model-specific notation
$\Phi(Q)$ = incremental cost savings function, which is the difference between conventional and flexible transit
$V_0(Q)$ = option value of using flexible bus service
$V_1(Q)$ = option value of using conventional bus service
F^+ = the cost of switching from the flexible bus service to the fixed bus service
F^- = the cost of switching from the fixed bus service to the flexible bus service
Q_H = upper demand trigger when switching from flexible transit to fixed transit
Q_L = lower demand trigger when switching from fixed transit to flexible transit
$H(\cdot)$ = the confluent hypergeometric function or Kummer function
F_s = switching cost from one vehicle size to the other vehicle size.

The approach taken for entry-exit problems is "asset equilibrium pricing". In other words, we construct differential equations for dynamic equilibrium values of the two different modes based on the stochastic process. If there exist thresholds between the two modes, then the parameters of those equations can be determined such that the state conditions and relations between the two value functions at those thresholds hold.

Let $V_0(Q)$ be the option value of using flexible bus operating mode when demand density is Q, and $V_1(Q)$ be the option value of using conventional bus operating mode. By Ito's lemma, the asset equilibrium condition is equivalent to the second order differential equation (Dixit, 1989) for flexible bus mode in Eq. (21).

$$\frac{1}{2}\sigma^2 Q^2 V_0''(Q) + \mu(m-Q)V_0'(Q) - \rho V_0(Q) = 0 \tag{21}$$

The option value for conventional bus service can be calculated similarly. The only difference is that there is additional immediate cost $\Phi(Q)$ relative to flexible transit operation, defined from Eq. (20). The value of the asset $V_1(Q)$ must satisfy the ordinary differential equation in Eq. (22).

$$\frac{1}{2}\sigma^2 Q^2 V_1''(Q) + \mu(m-Q)V_1'(Q) - \rho V_1(Q) + \Phi(Q) = 0 \tag{22}$$

In addition to the asset equilibrium conditions, the value functions must be related to each other. F^+ is the cost assumed for switching from flexible bus service to conventional bus service, while F^- is the cost of switching from conventional bus service to flexible bus service. Due to the presence of switching costs, there is not a single threshold to cross from one service to the other. Instead, the switching costs create two thresholds for switching (Dixit, 1989), which we denote by demand thresholds Q_L (from fixed to flexible) and Q_H (from flexible to fixed). When the switching costs approach zero, the thresholds should converge to a single value. These thresholds establish the following "value matching" relations between the option values in Eq. (23) and (24).

$$\mathrm{V}_0(Q_H) = V_1(Q_H) - F^+ \tag{23}$$

$$\mathrm{V}_1(Q_L) = V_0(Q_L) - F^- \tag{24}$$

and "smooth pasting" (i.e. required alignment of the first order conditions) in Eq. (25) and (26).

$$\mathrm{V}_0'(Q_H) = V_1'(Q_H) \tag{25}$$

$$\mathrm{V}_0'(Q_L) = V_1'(Q_L) \tag{26}$$

When demand density Q is assumed to grow as a mean reverting process as in Eq. (1), we have a general solution of $V_0(Q)$ using infinite series (Tsekrekos, 2010), as shown in Eq. (27):

$$V_0(Q) = \left[A_0 H(-\gamma_0, w_0, x) + B_0\left(\frac{2\mu m}{\sigma^2 Q}\right)^{1-w_0} H(1-\gamma_0-w_0, 2-w_0, x)\right] Q^{\gamma_0} \tag{27}$$

where A_0 and B_0 are arbitrary constants, γ_0 is the positive root of the quadratic equation in Eq. (28), and w_0 and x are shown in Eq. (29) and (30), respectively.

$$-\gamma^2 + \left(1 + \frac{2\mu}{\sigma^2}\right)\gamma + \frac{2\rho}{\sigma^2} = 0 \tag{28}$$

$$w_0 = 2 - 2\gamma_0 + \frac{2\mu}{\sigma^2} \tag{29}$$

$$x = \frac{2\mu m}{\sigma^2 Q} \tag{30}$$

$H(\cdot)$ is the confluent hypergeometric function or Kummer function, given by the following series representation (Slater, 1960) in Eq. (31).

$$H(\gamma, w, x) = 1 + \frac{\gamma}{w}x + \frac{\gamma(\gamma+1)x^2}{w(w+1)2!} + \frac{\gamma(\gamma+1)(\gamma+2)x^3}{w(w+1)(w+2)3!} + \cdots \tag{31}$$

The function has an asymptotic expression. When the demand $Q \to 0$, then $x \to \infty$ and V_0 should approach zero. Therefore, we have the asymptotic behavior of the Kummer function (Sødal et al., 2008) in Eq. (32).

$$\lim_{x\to\infty} H(\gamma, w, x) = \frac{\Gamma(w)}{\Gamma(\gamma)} e^x x^{\gamma - w} \tag{32}$$

where $\Gamma(\cdot)$ is the Gamma function. Thus we obtain a relation between A_0 and B_0 in Eq. (33).

$$\mathrm{B}_0 = -\frac{\Gamma(1-\gamma_0-w_0)\Gamma(\mathrm{w}_0)}{\Gamma(-\gamma_0)\Gamma(2-\mathrm{w}_0)} A_0 = \Delta_0 A_0 \tag{33}$$

Similarly, for $V_1(Q)$ the solution for Eq. (22) is obtained in Eq. (34):

$$\begin{aligned} V_1\big(Q(t)\big) = {} & \left[A_1 H(-\gamma_1, w_1, x) + B_1 \left(\frac{2\mu m}{\sigma^2 Q}\right)^{1-w_1} H(1-\gamma_1-w_1, 2-w_1, x)\right] Q^{\gamma_1} \\ & + E_t\left[\int_t^\infty \Phi\big(Q(s)\big) e^{-\rho(s-t)} ds\right] \end{aligned} \tag{34}$$

where A_1 and B_1 are arbitrary constants, γ_1 is the negative root of the quadratic Eq. (28), and w_1 is determined from Eq. (35).

$$w_1 = 2 - 2\gamma_1 + \frac{2\mu}{\sigma^2} \tag{35}$$

When demand $Q \to \infty$, the option value of conventional bus service $V_1(Q)$ must be solved subject to the boundary condition in Eq. (36).

$$\lim_{Q\to\infty} V_1\big(Q(t)\big) = E_t\left[\int_t^\infty \Phi\big(Q(s)\big) e^{-\rho(s-t)} ds\right] \tag{36}$$

When demand $Q \to \infty$, $\frac{2\mu m}{\sigma^2 Q} \to 0$, and the Kummer function in Eq. (34) becomes $H(-\gamma_1, w_1, x) \to 1$ due to Eq. (31). Therefore, Eq. (34) reduces to the following relation between A_1 and B_1 in Eq. (37).

$$\lim_{Q\to\infty}\left[A_1 + B_1\left(\frac{2\mu m}{\sigma^2 Q}\right)^{1-w_1}\right] Q^{\gamma_1} = 0 \tag{37}$$

Because $\gamma_1 < 0$, and $1 - w_1 < 0$ from Eq. (35), we end up with $B_1 = 0$. For specific choices of μ, σ, and ρ, Eqs. (33) and (38) can be substituted into the value matching and smooth pasting conditions in Eq. (23)-(26) to establish four equations and four unknown variables: Q_L, Q_H, A_0, and A_1. It can be shown that the optimal solution $\boldsymbol{X} = [Q_H, Q_L, A_0, A_1]'$ is uniquely determined by solving the system of nonlinear equations in Eq. (38).

$$\mathbf{F}(\mathbf{X}) = \begin{bmatrix} A_0 H_0(Q_H) Q_H^{\gamma_0} + (\Delta_1 A_0 - A_1) H_1(Q_H) Q_H^{\gamma_1} - E_t\left[\int_t^\infty \Phi(Q(s)|Q(t) = Q_H) e^{-\rho(s-t)} ds\right] + F^+ \\ A_0 H_0(Q_L) Q_L^{\gamma_0} + (\Delta_1 A_0 - A_1) H_1(Q_L) Q_L^{\gamma_1} - E_t\left[\int_t^\infty \Phi(Q(s)|Q(t) = Q_H) e^{-\rho(s-t)} ds\right] - F^- \\ A_0 M_0(Q_H) Q_H^{\gamma_0} + (\Delta_1 A_0 - A_1) M_1(Q_H) Q_H^{\gamma_1} + \dfrac{\partial E_t\left[\int_t^\infty \Phi(Q(s)|Q(t) = Q_H) e^{-\rho(s-t)} ds\right]}{\partial Q} \\ A_0 M_0(Q_L) Q_L^{\gamma_0} + (\Delta_1 A_0 - A_1) M_1(Q_L) Q_L^{\gamma_1} + \dfrac{\partial E_t\left[\int_t^\infty \Phi(Q(s)|Q(t) = Q_H) e^{-\rho(s-t)} ds\right]}{\partial Q} \end{bmatrix} \tag{38}$$

In addition, $H_i(Q) = H\left(-\gamma_i, w_i, \frac{\xi}{Q}\right)$, where $\xi = \frac{2\mu m}{\sigma^2}$, and $\Delta_1 = \Delta_0 \xi^{1-w_0}$. The variable M_i is determined from Eq. (39).

$$M_i(Q) = \gamma_i\left(H_i(Q) + \frac{\xi}{w_i Q}\left(1 - \gamma_i, 1 + w_i, \frac{\xi}{Q}\right)\right) \tag{39}$$

For any given set of parameter values, the system equation in Eq. (38) can be solved numerically. Since the functional form of $\Phi(Q)$ is a complex polynomial, the expression $E_t\left[\int_t^\infty \Phi(Q) e^{-\rho(s-t)} ds\right]$ is computed numerically using numerical integration, where the incremental value $P_i = \Phi(E[Q(s)]) e^{-\rho(s-t)}$ is computed in each of the N time slices, and $E_t\left[\int_t^\infty \Phi(Q) e^{-\rho(s-t)} ds\right] \approx \frac{1}{N}\sum_i P_i$.

Due to the complexity of $E_t\left[\int_t^\infty \Phi\big(Q(s)\big) e^{-\rho(s-t)} ds\right]$, we have to obtain its derivative numerically. We divide the demand into many small intervals ΔQ , and approximate the derivative by finite difference: $\frac{E_t\left[\int_t^\infty \Phi(Q(s)+\Delta Q) e^{-\rho(s-t)} ds\right] - E_t\left[\int_t^\infty \Phi(Q(s)-\Delta Q) e^{-\rho(s-t)} ds\right]}{2\Delta Q}$.

3.3. Model variation: modular vehicle size

The general structure presented in the earlier sections can be modified to deal with different variations. For example, instead of fixed versus flexible transit, the same model can be used to evaluate switching between two different vehicle sizes S_1 and S_2. In this variation, an operator runs modular vehicles that can operate in one of two different sizes: a single vehicle of size S_1 or a platoon of two vehicles operating as one of size $S_2 = 2S_1$. The model can determine when to switch from one size to the other across the system to best serve a random demand process.

There is only flexible bus service with no fixed transit service, and the vehicle size is exogenous to the cost function, i.e. $C_{sf1}(Q; S_1)$ pertains to the cost of operating flexible service with vehicle size S_1 and $C_{sf2}(Q; S_2)$ for operating flexible service with vehicle size $S_2 = 2S_1$. In this case, the incremental cost savings function variation Φ_v for the S_2 sized vehicle is shown in Eq. (40).

$$\Phi_\text{v}\big(Q(t)\big) = C^*_{sf1}(Q(t); S_1) - C^*_{sf2}(Q(t); S_2) \tag{40}$$

All the other equations remain the same, except that the flexible transit (0, f) subscripts are replaced by $f1$ and fixed transit (1, c) subscripts are replaced by $f2$. In the case of a modular fleet that can switch between vehicles sizes at negligible cost, the switching cost may be zero, leading to a single threshold between the two different operations.

4. Model properties

In this section, several properties of the model are illustrated to showcase its significance. For some effects, computational examples are needed due to the use of numerical integration to solve Eq. (38). For these examples, the following parameters in Table 2 are referenced unless stated otherwise. The demand density is set to a range from 30 to 40 trips/mi^2/hr. To put that in perspective, New York City had 1.76 billion annual subway trips, 651 million annual bus trips, and 8.83 million demand-responsive trips in 2015 (MTA, 2015). Assuming an area of 305 mi^2, annual-to-weekday factor of 1/311, and afternoon peak hour factor of 0.15, this leads to hourly trips in the range of 2791 subway trips, 1030 bus trips, and 14 demand-responsive transit trips per square mile-hr. The demand in the example falls in between DRT and fixed-route bus. The service is assumed to initiate with the fixed-route transit service, and vehicle sizes are dynamically optimized (i.e. Eqs. (11a) and (19a)). In Section 5, vehicle sizes are held fixed using Eqs. (11b) and (19b).

Table 2. Input parameters of the numerical example.

Variable	Baseline Value	Unit
a	100	($/bus hr)
b	0.1	($/seat hr)
d	0.2	(miles)
J	8	(miles)
L, W	5, 3	(miles)
u	1.2	-
V_c	20	(miles/hr)
V_f	15	(miles/hr)
V_x	2.5	(miles/hr)
v_i, v_w, v_x	5, 12, 12	($/passenger hr)
y_c, y_f	1.8, 2.0	-
z_c, z_f	1.8, 2.0	-
k	1.15	-
F^+	10	($)
F^-	10	($)
F_s	100	($)
m	40	(trips/mile^2/hr)
μ	0.2	-
σ	7	-
ρ	7%	-
$Q(0)$	32	(trips/mile^2/hr)

For a collection tour in a grid network, the k value is 1.15 (Daganzo, 1984a).

4.1. Sensitivity of switching policy to transportation system parameters

The value of the dynamic switching policy can be illustrated with an example involving a change in the local service speed. A decrease in speed can indicate increased road congestion, while an increase in speed can be due to improved technologies that reduce dwell times and such. The impact can be measured by taking the base conditions in Table 2 and adjusting the operating speeds V_c and V_f by $\pm 20\%$. The solutions to the option values and lower and upper demand thresholds for the switching policy are obtained by solving the system of equations in MATLAB.

In the base case illustrated in Table 3 ($\alpha = 1.0$), the total cost of the fixed and flexible transit services are nearly the same (2881.1-2883) and the optimal decision under the policy is to stay in fixed transit service because $Q(0) > Q_L$ and the initial operating service is fixed transit. However, as the speed changes from -20% to +20%, there is actually a change in the minimum cost service from fixed transit (-20%) to flexible transit (+20%). In the scenario $\alpha = 1.2$, it is optimal to switch to flexible transit because $Q(0) < Q_L$. This illustrates the sensitivity of the policy to the transit system conditions, and the asymmetry in the sensitivity of the Q_L and Q_H to those conditions.

A second example of the sensitivity can be illustrated with a change in the unit variable operating cost as shown in Table 4. A similar computation with changes in the parameter b shows that the total costs can fluctuate between fixed and flexible transit service. The baseline scenario ($\alpha = 1.0$) is shown alongside a scenario with $\pm 20\%$ reduction in operational variable cost b. This cost can represent the cost of fuel or driver wages, for example. In the baseline scenario, the optimal decision at time $t = 0$ is to stay in fixed transit service because $Q(0) > Q_L$. In the scenario where $\alpha = 0.8$, however, the optimal decision is to switch at time $t = 0$ to flexible transit service because $Q(0) < Q_L$.

Table 3. Effects of local service speed on the model outputs.

Solutions		Scaling factor α of basic value of V_c and V_f		
		$\alpha = 0.8$	$\alpha = 1.0$	$\alpha = 1.2$
Fixed Transit	Headway, h_c	0.43	0.42	0.39
	Vehicle size, S_c	80	75	66
	Fleet size, F_c	5	5	5
	Route spacing, r	1.46	1.41	1.33
	Total cost, C_{sc}	3044.0 (+5.65%)	2881.1	2626.1 (-8.85%)
Flexible Transit	Headway, h_f	0.08	0.08	0.07
	Vehicle size, S_f	8	7	7
	Fleet size, F_f	62	58	50
	Service zone, A	2.85	3.02	3.34
	Total cost, C_{sf}	3153.2 (+9.37%)	2883.0	2470.1 (-14.32%)
$\Phi(Q)$		109.2	1.9	-156
$E_t\left[\int_t^\infty \Phi(Q)e^{-\rho(s-t)}ds\right]$		654.8	-39.4	-989.0
$V_0(Q(0))$		16.5	23.5	125.8
$V_1(Q(0))$		35.9	17.4	76.7
Q_L		12.2 (-57.5%)	28.7	68.9 (+140.1%)
Q_H		21.5 (-48.6%)	41.8	83.1 (+98.8%)
Indifference band ($Q_H - Q_L$)		9.3 (-29.0%)	13.1	14.2 (+8.4%)

One of the features explored by Dixit (1989) was the notion of "hysteresis". The hysteresis effect due to the switching costs leads to an "indifference band" in the demand density that creates the lower and upper thresholds. In this case the width of the indifference band is 13.1 in the base case, but it changes under different transportation system scenarios, to as much as 14.2 and as little as 9.3 among the four alternative sensitivity scenarios.

Table 4. Sensitivity of variable bus operating cost on fixed and flexible transit total cost with baseline-relative changes in parentheses.

Solutions		Scaling factor α of basic value of operational variable cost b		
		$\alpha = 0.8$	$\alpha = 1.0$	$\alpha = 1.2$
Fixed Transit	Headway, h_c	0.42	0.42	0.42
	Vehicle size, S_c	75	75	75
	Fleet size, F_c	5	5	5
	Route spacing, r	1.41	1.41	1.41
	Total cost, C_{sc}	2873.1 (-0.28%)	2881.1	2888.9 (+0.27%)
Flexible Transit	Headway, h_f	0.08	0.08	0.08
	Vehicle size, S_f	8	7	7
	Fleet size, F_f	57	58	58
	Service zone, A	3.13	3.02	2.91
	Total cost, C_{sf}	2799.0 (-2.91%)	2883.0	2996.2 (+3.93%)
$\Phi(Q)$		-74.1	1.9	107.3
$E_t\left[\int_t^\infty \Phi(Q)e^{-\rho(s-t)}ds\right]$		-802.4	-39.4	995.5
$V_0(Q(0))$		86.4	23.5	19.8
$V_1(Q(0))$		44.5	17.4	33.5
Q_L		42.4 (+47.7%)	28.7	21.5 (-25.1%)
Q_H		55.6 (+33.0%)	41.8	31.2 (-25.4%)
Indifference band ($Q_H - Q_L$)		13.2 (+0.8%)	13.1	9.7 (-26.0%)

4.2. Switching policy sensitivity to demand density

The sensitivity of the policy to demand density is investigated further. The baseline parameters provided in Table 2 are used to plot the incremental cost savings function $\Phi(Q)$ with respect to Q using Eq. (20) in Fig. 4(a). This shows that, for the example, operating fixed transit would lead to a better immediate payoff if the demand density is 32 or higher. This is also the threshold that would be used in a deterministic setting with perfect information to determine switching points.

The option values V_0 (Eq. (27)) and V_1(Eq. (34)) are also plotted after solving for their parameters in the system of equations from the baseline data. The plot is shown in Fig. 4(b), where the option values cross at 35.6, but the optimal switching thresholds taking into account the switching costs would be at $Q_L = 28.7$ and $Q_H = 41.8$. Note that the value of 35.6 is above the incremental cost savings equilibrium point, which makes sense because the option value accounts for foresight. Since the future trend is for the demand to converge towards a mean value of 40, the option value cost-free threshold lies above the incremental threshold.

The indifference band can also be visualized in Fig. 4(b). Although the switching costs are symmetric (at a value of \$10) in this example, the gap between the crossover threshold and Q_L is 6.9, while the gap between Q_H and the crossover threshold is 6.2. This asymmetry is visually explained by the nonlinearity of the option values (V_0, V_1).

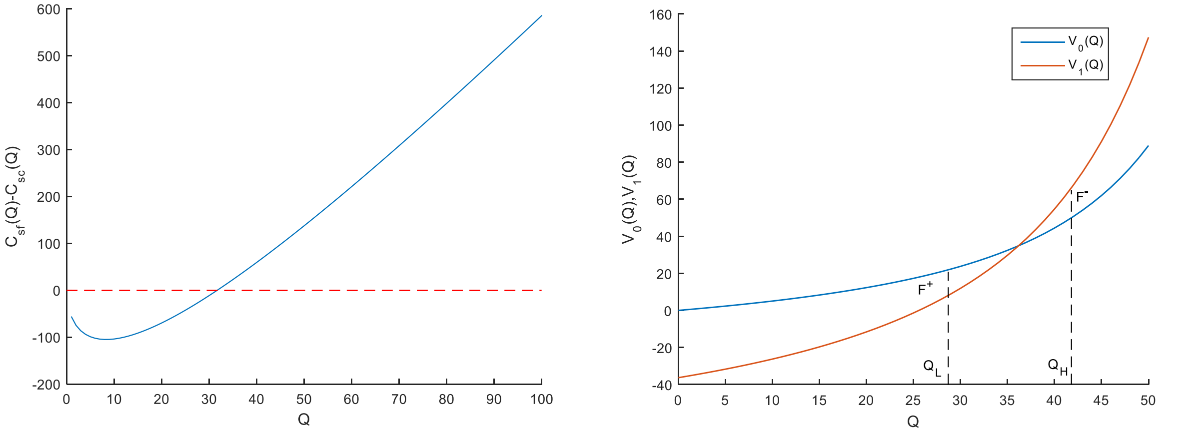

Fig. 4. (a) Incremental operational cost from fixed to flexible transit, (b) option value of conventional and flexible bus service with respect to demand density.

4.3. Switching policy value sensitivity to stochastic process parameters

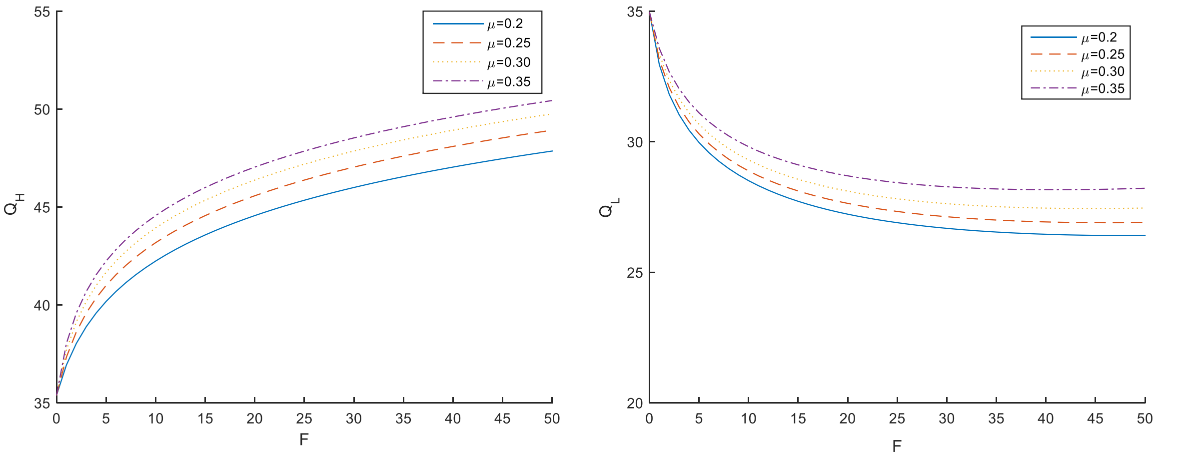

Fig. 5. Relation between (a) upper and (b) lower demand density threshold and switching cost for different reversion speed.

The prior two sections assessed the sensitivity to transportation system state and to initial demand. The policy is also dependent on the characteristics of the randomness in the demand, as measured directly using Fig. 5 to plot the policy thresholds under a range of reversion rates μ and switching costs F^+ and F^-.

From Fig. 5, with increasing switching cost from 0 to 50 ($/hr), the upper trigger demand density also increases, while the lower trigger decreases. This shows that the hysteresis effect emerges only when switching cost are present, and widens as those costs increase. Similarly, as the reversion rate increases, the upper threshold increases and the lower threshold decreases. One can conclude that when the switching costs increase, it is better to wait longer before intervening.

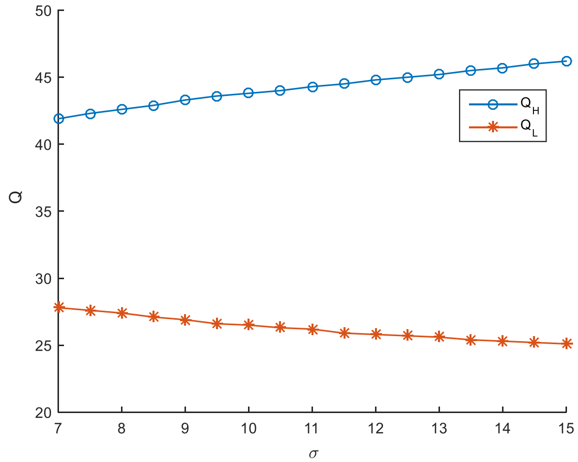

Fig. 6. Relation between upper and lower demand density threshold and volatility.

Fig. 6 shows the sensitivity of the upper and lower demand density threshold to the value of the volatility parameter σ. A rise in volatility raises the upper threshold Q_H and lowers the lower threshold Q_L. As the volatility increases, the indifference band widens. The intuition behind this result is that greater uncertainty would postpone invention.

Lastly, the discount rate is modified to examine the changes in the cost-free switching threshold and the upper and lower thresholds. These results are shown in Table 5. Just as in the earlier instances, the distances from the cost-free single threshold are not symmetric. Lastly, as the discount rate increases, we can see the crossover threshold moving up. This suggests that the present value of the flexible transit option value decreases at a slower rate than the fixed transit option value, leading to a shift upwards in the threshold.

Table 5. Relation between discount rate and trigger demand density thresholds.

Discount rate ρ (%)	Single trigger demand density Q^*	Upper demand density Q_H (trips/mile2/hr)	Lower demand density Q_L (trips/mile2/hr)
5	35.2	41.0 (+5.8)	28.3 (-6.9)
7	35.6	41.8 (+6.2)	28.7 (-6.9)
9	36.1	42.6 (+6.5)	29.1 (-7.0)
11	36.7	43.4 (+6.7)	29.5 (-7.2)

5. Computational evaluation of proposed policy

Two sets of numerical examples are shown here. The first set uses the same benchmark parameters set in Chang and Schonfeld (1991a) to show how the problem changes in a stochastic dynamic setting with certain stochastic process parameters. The objective is to evaluate for a sample trajectory how well the proposed policy performs relative to two other benchmark policies: a "perfect information" policy that assigns deterministically as if the operator knew the demand outcome beforehand; and a myopic policy that assigns whenever the incremental cost threshold ($\Phi = 0$) is crossed, as in Kim and Schonfeld (2015). The second set of examples is conducted to illustrate the value of autonomous vehicle modularity. Using the same modular vehicles provided by NEXT ($S_1 = 10$ passenger capacity), the value of switching between single vehicles and two-vehicle platoons will be quantified and compared against a system that operates under a fixed, optimal vehicle size for the average demand density.

5.1. Verification of policy against perfect information and myopic policies

5.1.1. Experimental design

The objective of this experiment is to compare the performance of the proposed policy against an ideal "perfect information" set of decisions and a myopic policy. The performance measure is the cumulative cost accrued over a simulated trajectory. Since the underlying stochastic process is stationary, it is not necessary to run multiple trajectories to obtain an average; a sufficiently long trajectory works just as well.

For the same simulated demand density trajectory, each policy's outcome decisions are made as follows in three scenarios:

- Perfect information scenario: determine the optimal switching points deterministically in similar fashion to the example in Section 2.1;
- Myopic policy scenario: simulate the dynamic evolution of the demand density, and choose to switch to fixed transit whenever $\Phi > 0$ and to flexible transit otherwise;
- Proposed policy scenario based on market entry-exit switching option: simulate the dynamic evolution of the demand density, and switch to fixed transit whenever $Q(t) > Q_H$ or switch to flexible transit whenever $Q(t) < Q_L$.

The duration of the trajectory is over 96 units of time, where each unit is assumed to be 15 minutes (equivalent to 24 hours). The parameters of the underlying O-U process are specified as follows: $m = 40$, $\mu = 0.2$, $\sigma = 7$, $\rho = 7\%$, $Q(0) = 32$, and the initial service is assumed to be flexible transit ($a(-1) = 0$). Here the vehicle sizes for fixed and flexible transit are fixed at 80 and 8 seats/vehicle, respectively, based on finding optimal vehicle sizes under average demand density of 40 trips/mile^2/hr. Simulation of the O-U process is performed by discretizing Eq. (1) and approximating it with Eq. (41):

$$Q_{n+1} = Q_n + \mu(m - Q_n)\Delta t + \sigma Q_n \Delta w_n \tag{41}$$

where Δw_n are independent identically distributed Wiener increments, i.e., normal variates with zero mean and variance Δt. Thus, $w_{t_{n+1}} - w_{t_n} = \Delta w_n \sim N(0, \Delta t) = \sqrt{\Delta t} N(0,1)$. $N(0,1)$ is the standard normal variate. In implementation, $\Delta t = 0.5$.

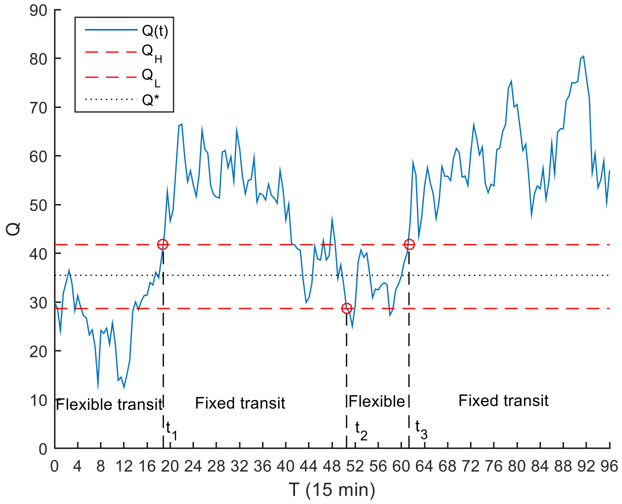

Fig. 7. Simulated mode switches from proposed switching policy.

Two types of costs should be accumulated in each scenario: cost of switching, and cost of operating with one type of transit service at the realized demand density. For performance measurement, the perfect information scenario should result in the lowest realized total cost over the 24-hour period (denoted R_{ph}), while the myopic is expected to operate the worst (denoted R_{my}). The cost accrued from the proposed policy is set to R^* . The performance of the proposed policy, $\varpi(\pi)$, is set to be a fraction where a value closer to 1 indicates a performance closer to the perfect hindsight relative to the myopic policy, as noted in Eq. (42).

$$\varpi(\pi) = \frac{R_{my} - \pi}{R_{my} - R_{ph}} \tag{42}$$

5.1.2. Results

Fig. 7 shows one simulation path of demand density over time overlaid with the thresholds for the proposed policy. When switching cost is 0, there is just one threshold, $Q^* = 35.6$. When there is a switching cost between conventional and flexible bus service of $F^+ = F^- = 10$, for example, then the optimal thresholds are $Q_L = 28.7$ and $Q_H = 41.8$. Under the policy, from 0 to t_1, and t_2 to t_3, the flexible bus service is selected; from t_1 to t_2, and after t_3, the fixed bus service is selected.

Fig. 8(a) illustrates the myopic policy; it shows the discounted incremental cost saving $\left(\Phi\left(Q(t)\right)e^{-\rho t}\right)$ over time with the demand density trajectory. Whenever the incremental total cost saving crosses zero (dotted line), it would switch. This is what would be followed according to the fixed/flexible service selection studies conducted without considering look-ahead using the dynamic demand information.

For the perfect information scenario, we use the mypic policy as a starting point, but remove any switches where the switching costs exceed the cumulative discounted cost saving. This leads to only one switching point, as shown in Fig. 8(b), assuming flexible transit service is the initial service.

The performance of our proposed policy is computed and shown in Table 6. The result suggests that, relative to the myopic policy, the proposed policy can reduce the excess cost by 72% in this 24-hour simulated example.

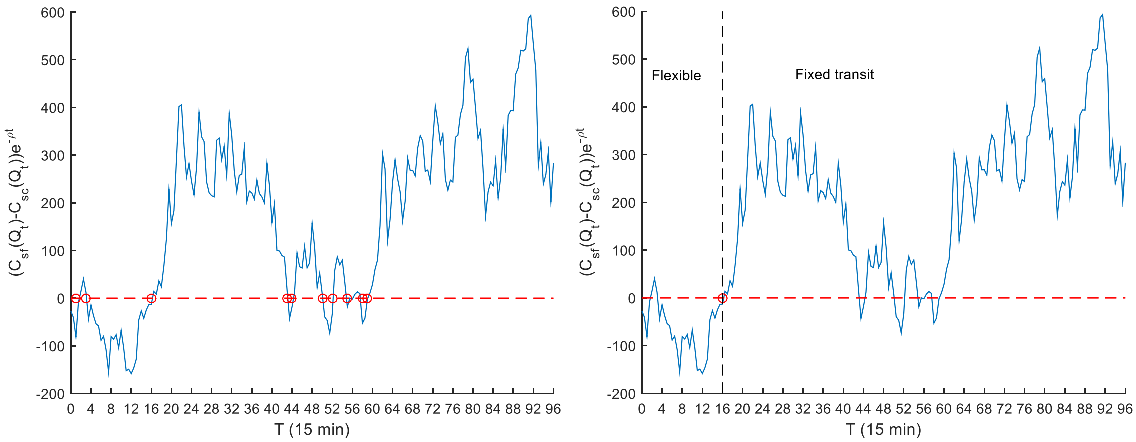

Fig. 8. (a) Myopic switching policy, and (b) perfect information switching policy.

Table 6. Performance comparison of the three policies.

	Perfect hindsight	Proposed policy	Myopic policy
Total discounted cost	46093.77	46150.62	46293.73
$\varpi(\pi)$	1.0000	0.7157	0.0000

5.2. Application to vehicle modularity

5.2.1. Experimental design

The case for having vehicle modularity is demonstrated. The experiment is designed to compute the option premium for the added flexibility to switch between two vehicle sizes: $S_1 = 10$ and $S_2 = 20$. In a modular fleet as illustrated in Fig. 1, the second size would be achieved by joining two vehicles to become a single vehicle. How much value does that add to an operation compared to having flexible service with vehicles of only one size?

The premium is tested using a computational example, taking the same demand density trajectory from Section 5.1. The difference is that the service operates under two flexible service scenarios:

- Scenario 1: flexible service with only one fixed vehicle size S_0 (static policy),
- Scenario 2: flexible service with two vehicle sizes in which the proposed policy is used to determine optimal switching, assuming the system initiates at S_1 and having symmetric switching costs $F_S = 10$.

The difference in accumulated cost, divided by the length of time (96 units), yields the premium of having the flexible vehicle sizing per unit time. The static vehicle size is set by finding the optimal vehicle size (Eq. (16)) for $Q = 40$ (since $m = 40$). This size is $S_0 = 8$.

5.2.2. Results

Based on the parameters, the total incremental costs with respect to the demand density and their differences between the two vehicle sizes (Φ_v from Eq. (40)) are plotted in Fig. 9. The demand where the incremental cost threshold lies is 25.5 trips/mi^2/hr.

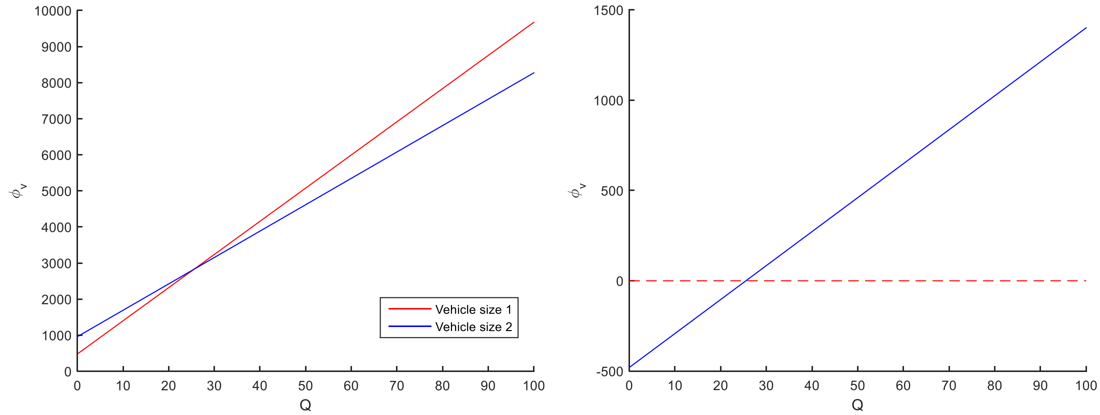

Fig. 9. (a) Total cost of two vehicle sizes and (b) cost savings of vehicle size 2 over vehicle size 1, with respect to demand density.

This cost saving function is then used to determine the policy decisions in the same simulated demand density trajectory shown in Fig. 10.

In the static policy where vehicle size is kept at $S_0 = 8$ throughout the 24 hours, the cumulative discounted cost is found by numerically integrating the cost function.

For the proposed policy, the solution is found to be an upper threshold of $Q_H = 33.5$ to switch from S_1 to S_2, and a lower threhold of $Q_L = 22.8$ to switch from S_2 to S_1, assuming the symmetric switching cost F_S at 10. When $F_S = 0$, the thresholds collapse into a single value $Q^* = 28.2$. The proposed policy leads to three switches shown in Fig. 10.

The total cumulative discounted costs are shown in Table 7. The flexibility to switch vehicle size in this case leads to an improvement over a static policy of \$373.45 over the 96 periods, i.e. a premium of \$3.89 per period.

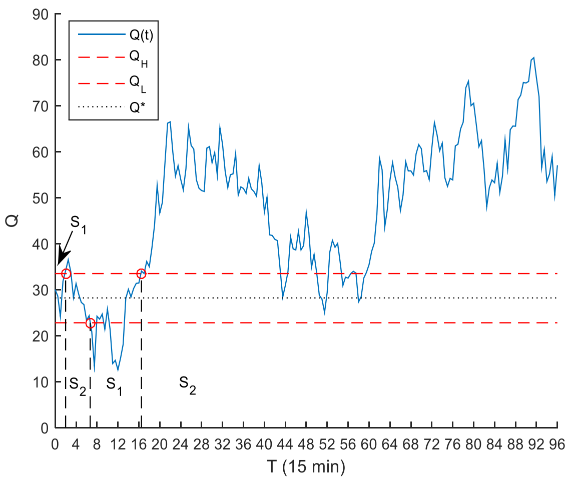

Fig 10. Proposed switching policy for vehicle modularity.

Table 7. Cumulative cost comparison of static policy and our policy.

	Proposed policy	Static policy
Cumulative total cost ($)	45946.64	46320.09

6. Conclusion

With the rise of many information-driven “smart transit” services (Sayarshad and Chow, 2015), there is a greater need than ever to use that information to guide dynamic decision-making. This work overcomes a major hurdle in the flexible transit service analysis literature by proposing the first anaytical model of dynamic service selection under time-variant uncertainty. Several contributions are made:

- We formulate the first analytical decision model to optimize dynamic switching of transit service as a market entry-exit real options model with mean-reverting demand density;
- We present a variation of this model to address “vehicle modularity”, which will be an important feature of shared autonomous vehicle fleets;
- By comparing against a benchmark computational example from Chang and Schonfeld (1991a), we show that:
  - There exists a hysteresis effect where there is a cost of switching from one service to another, and this effect is sensitive to transportation system conditions and demand characteristics;

- The values and sensitivity of the switching thresholds with respect to the transportation system conditions are not necessarily symmetric;
- The cost-free switching threshold is not equivalent to the deterministic cost savings threshold because the former accounts for look-ahead whereas the deterministic cost savings threshold is myopic;

- We validate the performance of the proposed policy by running it over a 24-hour period, and show that relative to a myopic policy from the prior literature, the proposed policy can reduce up to 72% of the excess cost;
- We show the existence of an option premium for having the flexibility to switch between two vehicle sizes.

The results are promising and can be taken in several different directions. The model can be used directly by current transit agencies to determine data-driven transitions between fixed and flexible last-mile transit services to reduce costs. When the Ornstein–Uhlenbeck process is changed to a Geometric Brownian Motion process, the model may be applicable to long term regional planning and staging of flexible-route and fixed-route transit services. The vehicle size switching variation will be of use to shared autonomous vehicle fleets such as NEXT. While the timing and control method has been applied to flexible and fixed route transit mode switching because of the availability of comparable benchmarks, the model framework can also be used to control mode-switching in other transportation systems (e.g. dynamic traffic or parking pricing, dynamic changeable lane directions or usage type (e.g. HOV), managing express service, dynamic vehicle routing, dynamic server relocation), with a little tweaking. However, for cases involving network effects, may need numerical methods (Chow and Regan, 2011a; Hinz and Yap, 2016).

Further research should also be undertaken to address some shortcomings or simplifications assumed in this study. The base transit service cost models assumed only a single many-to-one region for simplicity. Other service patterns based on the analytical model from Chang and Schonfeld (1991a) can also be considered that would handle concurrent mixed service, although optimal allocation of resources within the geographic areas and demand subgroups are not guaranteed. Hybrid analytical options valuation with computational network optimization may be used to reflect concurrent mixed operations. If travel speed or some other system parameter is stochastic, more advanced option valuation methods will be needed. Service switching is assumed to be nearly instantaneous relative to the time horizon in this study. Adding significant switching duration as a research extension should be feasible, given similar considerations of construction duration in other transit real options studies (e.g. Li et al, 2015). An empirical study using this methodology on a real transit service with performance validation would be highly valuable. For example, this method may show that current demand density does not suffice to justify a certain transit technology investment (e.g. heavy rail, light rail, or bus), and prefer more flexible operations that exploit switching.

Acknowledgements

The first author's work was conducted at the University of Maryland as a visiting researcher, with acknowledged funding from the Guangdong Provincial Natural Science Foundation under grant No. 2016A030310223 and National Science Foundation for Young Scientists of China under grant No. 71601192. Dr. Chow was partially supported by NSF grant CMMI-1634973. The first author also benefited from discussions with Drs. Zhongfei Li and Shuming Chen.

References

Ackerman, E., 2016. nuTonomy to test world's first fully autonomous taxi service in Singapore this year. IEEE Spectrum, http://spectrum.ieee.org/cars-that-think/transportation/self-driving/nutonomy-to-launch-worlds-first-fully-autonomous-taxi-service-in-singapore-this-year, last accessed July 28, 2016.

Adebisi, O., Hurdle, V. F., 1982. Comparing fixed route and flexible route strategies for intra-urban bus transit. *Transportation Research Record.* 854, 37-43.

Cadenillas, A., Lakner, P., Pinedo, M., 2010. Optimal control of a mean-reverting inventory. *Operations Research*. 58(6), 1697-1710.

Chang, S.K., Schonfeld, P., 1991a. Optimization models for comparing conventional and subscription bus feeder services. *Transportation Science* 25(4), 281-298.

Chang, S.K., Schonfeld, P., 1991b. Integration of fixed and flexible route bus systems. *Transportation Research Record*. 1308, 51-57.

Chow, J.Y.J., Regan, A.C., 2011a. Real option pricing of network design investments. *Transportation Science* 45(1), 50-63.

Chow, J.Y.J., Regan, A.C., 2011b. Network-based real option models. *Transportation Research Part B* 45(4), 682-695.

Chow, J.Y.J., Regan, A.C., 2011c. Resource location and relocation models with rolling horizon forecasting for wildland fire planning. *INFOR* 49(1), 31-43.

Chow, J.Y.J., Regan, A.C., Ranaiefar, F., Arkhipov, D.I., 2011. A network option portfolio management framework for adaptive transportation planning. *Transportation Research Part A* 45(8), 765-778.

Chow, J.Y.J., Sayarshad, H.R., 2015. Reference policies for non-myopic sequential network design and timing problems. *Networks and Spatial Economics*, in press, doi: 10.1007/s11067-015-9315-5.

Daganzo, C. F., 1984a. The length of tours in zones of different shapes. *Transportation Research Part B* 18(2), 135-145.

Daganzo, C. F., 1984b. Checkpoint dial-a-ride systems, *Transportation Research Part B* 18(4), 318-327.

Dixit, A.K., 1989. Entry and exit decisions under uncertainty. *Journal of Political Economy* 97(3), 620-638.

Dixit, A.K., Pindyck, R.S., 1994. *Investment Under Uncertainty*. Princeton University Press, Princeton, NJ.

Ditlevsen, S., Ditlevsen, O., 2008. Parameter estimation from observations of first-passage times of the Ornstein–Uhlenbeck process and the Feller process. *Probabilistic Engineering Mechanics*, 23(2), 170-179.

Djavadian, S., Chow, J.Y.J., 2016. Agent-based day-to-day adjustment process to evaluate dynamic flexible transport service policies. *Transportmetrica B*, in press, doi: 10.1080/21680566.2016.1190674.

Gkochari, C. C., 2015. Optimal investment timing in the dry bulk shipping sector. *Transportation Research Part E*. 79, 102-109.

Fu, L., Ishkhanov, G., 2004. Fleet size and mix optimization for paratransit services. *Transportation Research Record*, 1884, 39-46.

Hinz, J., & Yap, N. (2016). Algorithms for optimal control of stochastic switching systems. *Theory of Probability & Its Applications*, 60(4), 580-603.

Hu, Y., Nualart, D. 2010. Parameter estimation for fractional Ornstein–Uhlenbeck processes. Statistics & Probability Letters, 80(11), 1030-1038.

Jacobson, J., 1980. Analytical models for comparison of alternative service options for the transportation handicapped. *Transportation Research*, 14B, 113-118.

Kim, M., Schonfeld, P., 2012. Conventional, flexible and variable-type bus services. *Journal of Transportation Engineering*, ASCE, 138(3),. 263-273.

Kim, M., Schonfeld, P., 2013. Integrating bus services with mixed fleets. *Transportation Research Part B* 55, 227-244.

Kim, M., Schonfeld, P., 2014. Integration of conventional and flexible bus services with timed transfers. *Transportation Research Part B* 68, 76-97.

Kim, M., Schonfeld, P., 2015. Maximizing net benefits for conventional and flexible bus services. *Transportation Research Part A* 80, 116-133.

Lee, K. T., Kuo, S. H. F., Schonfeld, P., 1995. Optimal mixed bus fleet for urban operation. *Transportation Research Record*, 1503 39-48.

Li, X., Quadrifoglio, L., 2010. Feeder transit services: choosing between fixed and demand responsive policy. *Transportation Research Part C* 18(5), 770-780.

Li, Z.C., Guo, Q.W., Lam, W.H.K., Wong, S.C., 2015. Transit technology investment and selection under urban population volatility: a real option perspective. *Transportation Research Part B* 78, 318-340.

Metcalf, G.E., Hassett, K.A., 1995. Investment under alternative return assumptions: comparing random walks and mean reversion. *Journal of Economic Dynamics and Control* 19(8). 1471-1488.

MTA, 2015. MTA 2015 Annual Report, http://web.mta.info/mta/compliance/pdf/2015_annual/2015_Narrative.pdf.

Oksendal, B., 2013. Stochastic differential equations: an introduction with applications. Springer Science & Business Media.

Parpas, P., Webster, M., 2014. A stochastic multiscale model for electricity generation capacity expansion. *European Journal of Operational Research*, 232(2), 359-374.

Powell, W.B., 2011. Approximate dynamic programming: solving the curses of dimensionality, 2nd edn. John Wiley and Sons, New York.

Psaraftis, H.N., Wen, M., Kontovas, C.A., 2016. Dynamic vehicle routing problems: three decades and counting. *Networks* 67(1), 3-31.

Quadrifoglio, L., Li, X., 2009. A methodology to derive the critical demand density for designing and operating feeder transit services. *Transportation Research Part B* 43, 922-935.

Qiu, F., Li, W., Haghani, A., 2015. A methodology for choosing between fixed-route and flex-route policies for transit services. *Journal of Advanced Transportation* 49(3), 496-509.

Sarkar,S., 2003. The effect of mean reversion on investment under uncertainty. *Journal of Economic Dynamics and Control* 28(2), 377-396.

Saltzman, A., 1973. Para-transit: taking the mass out of mass transit. *Technology Review* 75(8), 46-53.

Sayarshad, H.R., Chow, J.Y.J., 2015. A scalable non-myopic dynamic dial-a-ride and pricing problem. *Transportation Research Part B* 81(2), 539-554.

Slater, L. J. 1960. Confluent hypergeometric function. Cambridge University Press.

Sødal, S., Koekebakker, S., Adland, R., 2009. Value based trading of real assets in shipping under stochastic freight rates. *Applied Economics* 41 (22), 2793–2807.

Sødal, S., Koekebakker, S., Aadland, R., 2008. Market switching in shipping—a real option model applied to the valuation of combination carriers. *Review of Financial Economics* 17(3), 183-203.

Spera, E., 2016. CAREEM to bring driverless transportation solutions to the MENA region through partnership with NEXT Future Transportation. Linkedin, https://www.linkedin.com/pulse/careem-bring-driverless-transportation-solutions-mena-emmanuele-spera, last accessed July 28, 2016.

Stein, D.M., 1978. Scheduling dial-a-ride transportation systems. *Transportation Science* 12(3), 232-249.

Thomas, B.W., 2007. Waiting strategies for anticipating service requests from known customer locations. *Transportation Science* 41(3), 319-331.

Tsekrekos, A.E., 2010. The effect of mean reversion on entry and exit decisions under uncertainty. *Journal of Economic Dynamics & Control* 34(4), 725-742.

Tsekrekos, A.E., Yannacopoulos, A.N., 2016. Optimal switching decisions under stochastic volatility with fast mean reversion. *European Journal of Operational Research* 251(1), 148-157.

Yuan, J., Zheng, Y., Zhang, L., Xie, X., Sun, G., 2011. Where to find my next passenger? Proc. UbiComp 2011, New York, NY.

Zagoraiou, M., & Baldi Antognini, A. 2009. Optimal designs for parameter estimation of the Ornstein–Uhlenbeck process. Applied Stochastic Models in Business and Industry, 25(5), 583-600.