Bus headway optimization with consumer surplus as a measure of societal benefit

Jason Hawkins a[1], S.C. Wirasinghe a, J.D. Hunt a

a University of Calgary, 2500 University Drive NW, Calgary, Alberta, T2N 1N4, Canada

Abstract

This paper derives the square root bus dispatch optimization developed by Newell, using consumer surplus maximization, and extends it to consider dispatch-variant demand. Results are tested for City of Edmonton mode choice and ridership. The sensitivity of the function to wait time cost and passenger demand is tested. This suggests, for observed values of time and passenger demand, there is a portion of riders who are captive and will suffer increasing wait time costs, but most passengers will make the choice to utilize other modes of travel. It is determined consumer surplus provides a clearer optimal headway for a relatively inelastic logit-based demand function than the wait time cost used by Newell, but the two are equivalent for totally inelastic demand.

Keywords: bus dispatch; headway optimization; consumer surplus; logit model.

Acknowledgments
Thank you to Alan Brownlee for the valuable contributions of data and insight. This research was supported by the University of Calgary, NSERC, and a Queen Elizabeth II Scholarship.

1 Introduction

In bus operations, the dispatch rate for a given route is the rate at which vehicles are deployed to the route, in units of vehicles per unit time. It is one of the key elements of service design selected by the operator. If reliability of bus arrivals along the route is maintained at a high level, for example through the use of time points (Wirasinghe and Liu 1995; Vuchic 2007), then for users at points along the route (a) the expected service frequency, in units of vehicle arrivals per unit time, is equal to the dispatch rate, and (b) the mean headway, or expected time between vehicle arrivals, in units of time per vehicle, is equal to the inverse of the dispatch rate.

The dispatch rate impacts both operator costs and user costs. A higher dispatch rate requires more operating vehicles, which increases operator costs. A higher dispatch rate also reduces mean headways and thereby user wait times and thus user costs. The selection of a dispatch rate includes making a trade-off between its impacts on both operator and user costs. Wait time cost is a function of user value of time - typically dependent on income demographics – often derived from survey data (Hossain, Hunt, and Wirasinghe 2015).

Most previous studies have developed dispatching objective functions based on an assumption of total inelasticity of demand with headway. Intuitively, it can be said that as headways are decreased, demand for the route will increase due to increased attractiveness compared with other travel mode choices. Empirical support

[1]* Corresponding author. Tel.:+1-403-464-6909.
E-mail address: jfhawkin@ucalgary.ca

for this assertion is summarized in a TRL report, which finds the elasticity to be -0.26 across studies conducted in London, Dallas, and Dublin (Balcombe et al. 2004). An objective function for bus dispatch rate will be developed in this paper with user demand as a function of bus headway and Newell's “square root” policy (Newell 1971) as a foundation. Policy formulations, assuming both totally inelastic and relatively inelastic conditions, will then be applied to the City of Edmonton using a validated logit model.

2 Literature Review

The original idea of a square root relationship between demand and dispatch rate has been attributed to William Vickrey during a conversation with Herbert Mohring in the 1960s. Mohring discussed this theory in a paper on headways for a route with time invariant and uniformly distributed demand (Mohring 1972). A more sophisticated square root formula for the dispatch rate was developed by Newell (1971) ; there is no indication the two were aware of each other’s work. Newell’s initial analytical formulation is for a single route between an origin and destination with infinite capacity vehicles. He then expands this to analyse a many-to-one route: a route with several boarding locations, but one alighting point. Newell assumes a continuous curve for passenger arrival rate - p(t) – where demand is allowed to be time (t) variant and describes a peak period (implying commuters) served by buses with time variant headways.

The total cost function for a bus route per unit time, as it relates to headway, can be expressed as

$$z(t) = \left(\frac{1}{2}\right)\gamma_w p(t)h(t) + \frac{\lambda_D}{h(t)} \tag{1}$$

where z(t) is the total cost per unit time, γ_w is the mean unit wait time cost per passenger, and λ_D is the unit dispatch cost. The average wait time for a passenger is assumed to be half a bus headway. In the case of commuters who travel every week day, know the schedule, and arrive just in time to catch their chosen bus, this is a correct assumption, though the waiting is done at the destination. This is also known as schedule delay.

This gives a total cost minimizing headway

$$h^*(t) = \left(\frac{2\lambda_D}{\gamma_w p(t)}\right)^{0.5} \tag{2}$$

This policy formulation, commonly referred to as the Newell “square root” Policy, states that the optimal time variant headway h(t) for a large fleet of infinite capacity buses is inversely proportional to the square root of the passenger arrival rate. Newell (1971) identifies a constraint on this solution. He considers modification of dispatch rate with restriction on vehicle capacity. Vehicles will be dispatched when full, if they are filled prior to the square root headway being reached, resulting in

$$h_c(t) = \frac{c}{p(t)} \tag{3}$$

Thus

$$h^*(t) = \min \begin{cases} \left(\frac{2\lambda_D}{\gamma_W p(t)}\right)^{0.5} \\ \frac{c}{p(t)} \end{cases} \quad (4)$$

where dispatch rate is determined by assuming the square root law until arrival rate reaches a threshold at which passengers per bus equals vehicle capacity. At this point, buses are dispatched full at a rate inversely proportional to the capacity of the chosen vehicle. Newell (1971) also estimates the fleet size when buses are filled during a portion of a peak period and otherwise dispatched according to the square root law.

Wirasinghe (1990) expands the work of Newell to include many-to-many time-varying demand during a peak period served by buses with time variable headways. Adaptation of the Newell Policy to the many-to-many case can be accomplished by replacing the value of arrival rate, p(t), with demand estimated in terms of seats per unit time, s(t), for the purpose of estimating the capacity dispatch rate. This recognizes that a single seat can be occupied by multiple passengers in series as passengers will alight along the route, making their seat available to newly boarding passengers. The cumulative boarding and alighting rates can be compared and a maximum load value determined in order to calculate the capacity headway with respect to time. The maximum load point on a route is allowed to vary from dispatch to dispatch.

Wirasinghe (1990) further elaborates on the square root policy by examining a route that is served by uniform scheduled headways as opposed to time variant headways. This is analogous to replacing the time variable demand, p(t) in (4), with an average demand, $\bar{p}$, during a given time period. This simplification is adopted by a good portion of the transit literature (Hossain, Hunt, and Wirasinghe 2015). Further extensions by Wirasinghe (1990) include (i) considering a policy headway – a maximum allowable headway set by the transit agency – and (ii) headways based on considering the perceived cost of a unit of waiting time to be increasing linearly with elapsed wait time, which is applicable when demand is relatively low and headways tend to be larger, say >30 minutes.

Chang and Schonfeld (1993) complete a review of transit service design algorithms, which examines 13 examples. They find that 12 of the 13 studies assume a totally inelastic demand function, with demand uniformly distributed over the service area. Most of these algorithms also assume a many-to-one bus route and use an objective function that minimized operator cost and user wait time. Verbas and Mahmassani (2015) develop a network-level optimization using the Transit Network Frequency Setting Problem (TNFSP) formulation originally proposed by Furth and Wilson (1981). Two objectives are explored: 1) Maximizing the number of riders and the total waiting time savings under budget, fleet, policy headway and bus loading constraints. 2) Minimizing the net cost under fleet, policy headway, bus loading, minimum ridership and minimum waiting time savings constraints. They consider spatial and temporal heterogeneity of ridership with respect to headway. Verbas and Mahmassani (2015) apply this model to the Chicago Transit Authority system. This algorithm finds optimal solutions for both users and operators for both objective functions.

Kocur and Hendrickson (1982) develop a dispatch function assuming relatively inelastic demand for a grid network of local bus routes. They define consumer surplus as the total value of user willingness to pay based on a utility function, less the cost they pay in terms of fare and travel cost. They evaluate three objective function formulations: profit maximization, user benefit maximization under unconstrained conditions, and user benefit maximization constrained by vehicle size. They use a linear mode choice model to represent the relationship between headway and demand, prefaced on the fact that a typical logit demand function is more difficult to differentiate and manipulate. Their measure of consumer surplus is represented as

$$\left(-\frac{TpXY}{2\alpha_4}\right) * \left(\alpha_1 + \alpha_2\left(kh + \frac{g+b}{4j}\right) + \alpha_3\frac{d}{v} + \alpha_4 f + \alpha_5 d\right)^2 \left(\alpha_1 + \alpha_2\left(kh + \frac{g+b}{4j}\right) + \alpha_3\frac{d}{v} + \alpha_4 f + \alpha_5 d\right)^2 \tag{5}$$

where

$\alpha_1, \alpha_2, \alpha_3, \alpha_4,$ and α_5 are constant parameters, and
T = Time period of analysis (minutes)
p = Trip density by all modes (trips/mile^2/minute)
X = Width of analysis or service area (miles)
Y = Length of analysis area (miles)
k = Ratio of expected user wait time to headway
h = Headway on a local route (minutes)
g = Spacing between parallel bus routes (miles)
b = Spacing between bus stops along a route (miles)
j = Average walking speed (miles/minute)
d = Average passenger trip length (miles)
v = Average local bus speed, including stops (miles/minute)
f = Bus fare for local service (US cents)

This formulation measures consumer surplus as a function of trip density for all modes, service area, several parameters of wait time, spacing of bus routes and stops, bus fare, and average passenger trip length. It is a complex value to determine and has a high data requirement. The present formulation seeks to provide simplified metrics, with minimal loss of accuracy in representation of route characteristics.

3 Dependency of Demand Function on Bus Headway

3.1 Derivation from Edmonton, Alberta Data

A logit utility function for the City of Edmonton, Alberta was used to provide a representation of the relationship between user demand for buses on a corridor with changes in bus headway. The logit function was developed by the City of Edmonton based on calibration using the ALogit software package and a 2004 survey (City of Edmonton 2004). The following function for bus transit during the AM peak was used in the present analysis

$$V(bus) = a_1 t_B + a_2 t_a + \frac{a_3\, h(t)}{2} + a_4 T + a_5 F + a_6 \tag{6}$$

with variables as defined in Table 1.

To develop a relationship between demand and bus headway, several parameter estimates were required. Estimates for the time components of a standard bus trip (walk time, wait time, and ride time) and average number of transfers were derived from the EMME (Florian et al. 1979) results for the Edmonton Regional Travel Model (RTM). These estimates were based on model transportation skims for the current year and transit network. The National Household Travel Survey (Statistics Canada 2011) and a study conducted in Calgary, Alberta - the sister city to Edmonton - on walking distance to bus stops (Lam and Morrall 1981) were utilized to validate the model outputs and provided a mean auto commuter trip time of 23 minutes for Edmonton. The EMME estimate for total bus trip time matched the national average published by Statistics Canada within 0.5 minutes. The 2014 Edmonton Transit Service fare of $3.00 was used (City of Edmonton 2014). In the application to forecasting models, the City of Edmonton assumed the value of transit wait time is equal to one half a bus headway. By assuming all other variables constant, these assumptions represent the effect on demand of varying the headway for an average trip. Values used in the development of a representative curve of the relationship between headway and demand are summarized in Table 1.

Table 1. Summary of assumed variables in demand estimation.

Variable	Description	Value
t_B	Transit Ride Time	27.0 minutes
t_a	Transit Walk Time	8.6 minutes
h(t)	Bus Headway	variable
T	# Transfers per Trip	0.10
F	User Fare	$3.00
t_A	Auto Ride Time	23.0 minutes
a_1	Transit Ride Time Coefficient	-0.0597
a_2	Walk Time Coefficient	-0.0919
a_3	Wait Time Coefficient	-0.0919
a_4	Transfer Coefficient	-0.1858
a_5	Fare Coefficient	-0.5278
a_6	Transit Constant	3.5040
M	Fixed Component of Transit Mode Choice Representative Utility (Less Variable Wait Time Utility)	-0.5002
-	Fixed Bus Share for Totally Inelastic Case	21% (City of Calgary 2013)
γ_w	Passenger Wait Time Cost	$10.45/passenger-hour (City of Edmonton 2004)
λ_D	Bus Dispatch Cost	$80/bus (Calgary Transit 2002)

The relationship between the proportion of route users choosing to travel by bus can then be plotted against headway. Data were plotted from 0 minutes to 60 minutes using the utility function for the City of Edmonton (Fig. 1). This relationship results in diminishing returns on investment with increasing reductions of headway: a smaller change in consumer surplus for a unit change in user demand. We define consumer surplus as the total wait time for all users at the origin and/or destination, with changes in consumer surplus being a function of headway – wait time being half a bus headway on average. The marginal benefit will therefore approach the marginal cost, as represented by dispatch cost, and an optimal headway for maximization of consumer surplus can be derived according to the relationship illustrated in Fig. 4.

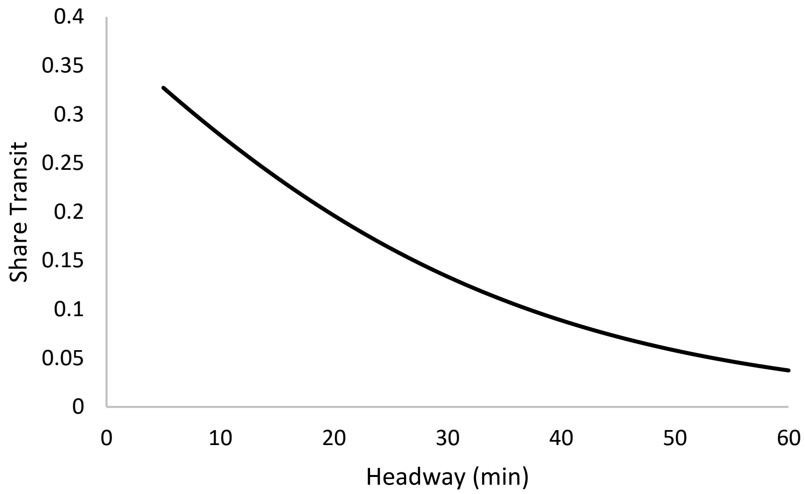

Fig. 1. Changing demand with variable headway (City of Edmonton 2004).

4 Methodological Framework

The square root dispatch policy developed by Newell (1971) assumes a totally inelastic relationship between the rate of dispatch of buses on a route and the passenger demand. This means, with an increase in dispatch rate, the demand on the route is assumed to remain constant. Wait time forms part of the cost included in the Newell Policy objective function. However, a higher dispatch rate will decrease the wait time of users. As the number of riders increases, the wait time each rider experiences decreases due to the need for higher dispatch rates. This effect is known as the Mohring Effect (1972) and is termed an "anti-congestion" effect (Mohring 1972) as an increase in the number of users lowers the cost of travel for existing users subject to capacity being available. This contrasts with the typical road congestion paradigm, wherein additional users create congestion and decrease the quality of service for existing users.

The methodology employed by Newell for a many-to-one route forms a sound theoretical foundation for development of a dispatch policy as it has a simple form and most inputs are fixed values. It can be readily extended by introducing a variable form for passenger arrival rate, demand. This basic formulation is modified with an assumption of relatively inelastic demand with respect to buses dispatched at uniform, scheduled, headways and buses never being filled to capacity. It is assumed that passengers are served by the first bus to arrive. The cost of a dispatch is assumed as constant with increasing dispatch rate. The change in the total cost of travel is therefore a function of the change in wait time multiplied by a unit wait time cost. The change in consumer surplus can be defined as

$$\Delta \mathrm{CS} = \frac{\Delta \mathrm{Headway}}{\Delta \mathrm{Passenger\ Arrival\ Rate}} \tag{7}$$

If passenger arrival rate is assumed as insensitive to changes in headway, as in the Newell Policy, the consumer surplus is as in Fig. 2. Under these conditions, total surplus is a rectangular integral of the change in headway at a constant passenger arrival rate of p_o between headways of h_1 and h_2. The total wait time cost savings are therefore solely a function of increases in vehicle dispatch rate.

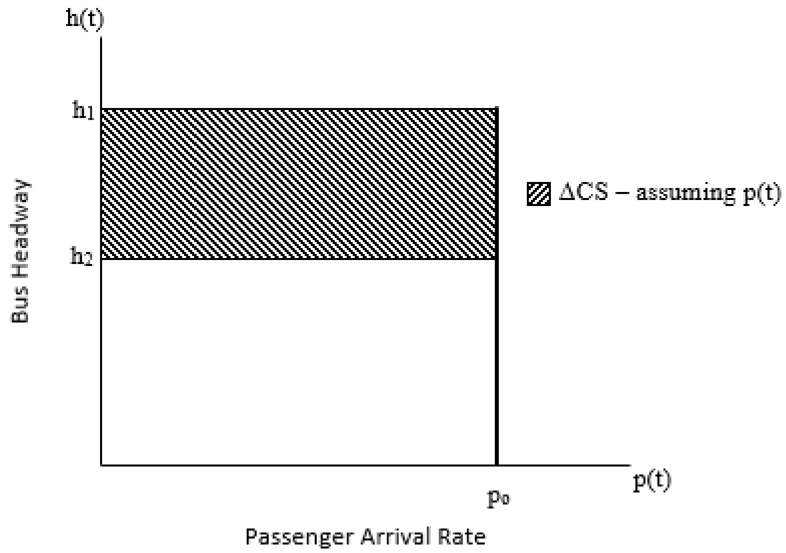

Fig. 2. Bus headway vs. passenger arrival rate – totally inelastic case.

By introducing demand elasticity with dispatch rate into the functional form, an additional consumer surplus is introduced from the resulting increase in passenger arrival rate (user demand). Fig. 3 illustrates this curve, with the additional consumer surplus due to elasticity of demand differentiated from the surplus under an assumption of total inelasticity. The consumer surplus for users shifting from other modes will be smaller as these users have a higher threshold for entry: they will not take the bus until the dispatch rate has risen above a certain, higher, threshold value.

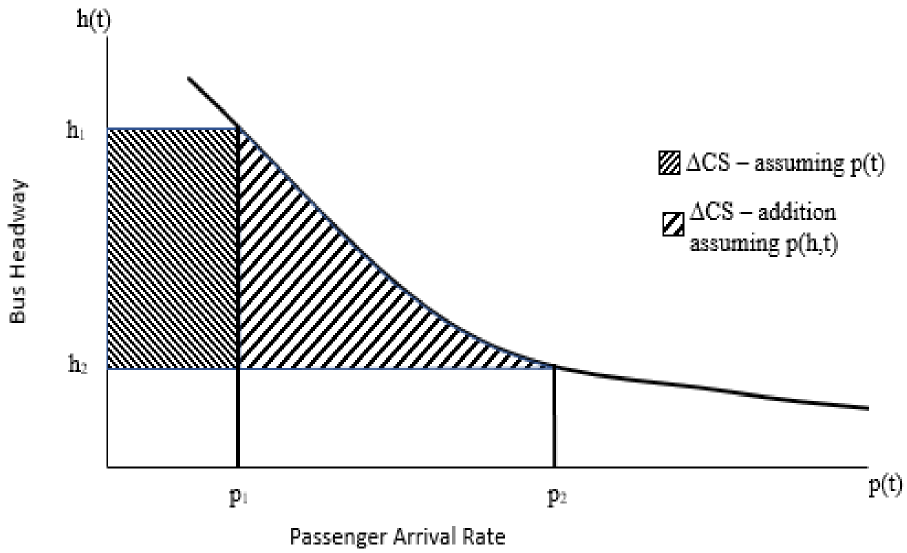

Fig. 3. Bus headway vs. passenger arrival rate – relatively inelastic case.

Fig. 3 suggests an additional surplus is accrued to users – not accounted for by the Newell Policy. This suggests that a change in headway from h_1 to h_2 would induce a corresponding change in passenger demand from p_1 to p_2 according to each person's elasticity to wait time – measured as a function of bus headway. An objective function was developed and the results compared with the Newell Policy under similar assumptions. These illustrations suggest headway could be reduced even to zero time, with consumer surplus continuing to increase in proportion to the demand function. One must therefore introduce a dispatch cost to determine the point where increasing dispatch rates result in negative marginal returns on investment. This is known as the point of "break-even marginal returns" (Fig. 4).

In the Newell Policy formulation, the optimal dispatch rate is determined by equating the derivative of the total cost function to zero. This gives the point at which an increase in dispatch rate does not result in a change in the objective cost. A similar condition is produced in the present formulation by comparing the change in consumer surplus to the dispatch cost. The marginal cost with respect to dispatch rate can be assumed constant as there is a constant cost associated with dispatching a bus. The marginal benefit will decrease as there is a non-linear increase in passenger demand for a unit increase in dispatch rate. The marginal wait time savings on a per passenger basis will therefore decrease with increasing dispatch rate. We can then define a net marginal return curve representing the combined marginal returns as a summation of wait time savings and additional dispatch cost. This net return will decrease with decreasing marginal benefit and reach a point of zero additional returns, the break-even marginal return, which is denoted by BEMR in Fig. 4. Below this point, the cost of dispatching an additional bus will outweigh the additional benefit to passengers resulting from wait time cost savings.

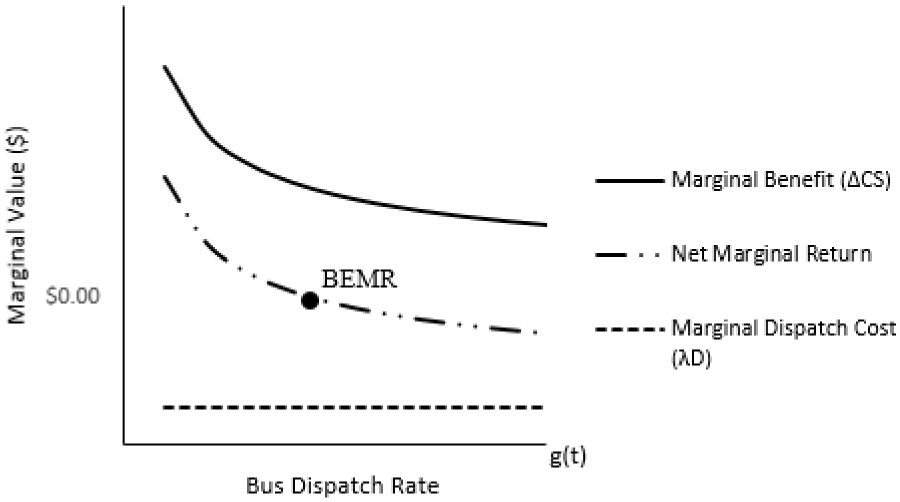

Fig. 4. Marginal return on changes in dispatch rate.

5 Optimization Function

Utilizing the same core cost function assumed as in Newell's Policy, cost per unit time, z(t), can be developed from the perspective of consumer surplus. Demand will be represented as a function of passenger wait time and a time varying arrival rate. Wait time is utilized as a measure of the level of transit service from the perspective of the passenger and can be approximated as half of a bus headway for small headways (Balcombe et al. 2004). It should be noted that for large headways, it is unlikely a passenger will wait at the stop for the full half of a headway, but this accounts for the difference in service between a 30-minute headway and 60-minute headway. This is termed the *schedule delay*, whereby the passenger will wait at home or work (Balcombe et al. 2004). In both cases the passenger may only wait 15 minutes at the bus stop, but the implicit wait time is assumed 30 minutes for the longer headway to account for the lower level-of-service associated with time spent waiting at the origin or destination. The two components of the objective function are

$$z(t)_{\text{Passenger}} = \gamma_w \int p(w,t)\,dh = \gamma_w \int p\left(\frac{h}{2},t\right) dh \quad (8)$$

$$z(t)_{Operator} = \frac{\lambda_D}{h} \quad (9)$$

A change in consumer surplus can then be defined as the change in wait time cost for a differential change in headway, h. The optimal dispatch rate will occur at the point of "break-even marginal return", that is where the change in consumer surplus equals the cost of dispatching (Fig. 4). This results in the following formulation

$$\frac{\partial CS}{\partial h} = \gamma_w \frac{\partial\left(\int \mathrm{p}\left(\frac{\mathrm{h}}{2},\mathrm{t}\right) dh\right)}{\partial h(t)} = -\lambda_D \frac{\partial(\frac{1}{h})}{\partial h} \tag{10}$$

$$\frac{\partial CS}{\partial h} = \frac{\gamma_w}{2}\mathrm{p(h,t)} = \lambda_D \frac{1}{h^2} \tag{11}$$

If it is assumed that passenger demand does not vary with changes in bus headway, that it is totally inelastic to headway and equal to p(t), the familiar formulation of Newell (1) emerges. The magnitude of the consumer surplus is then equivalent to the passenger wait time cost.

$$\mathrm{h}^* = \left(\frac{2\lambda_D}{\gamma_\mathrm{w}\mathrm{p(t)}}\right)^{0.5} \tag{12}$$

5.1 Logit Demand Function

The general form of the function presented in (11) can be applied to a variety of standard demand functions to introduce a deviation away from total inelasticity. In this paper, we utilize a logit representation as it is a commonly applied method and data were readily available. This will replace the value of p(h,t) with a logit-based approximation of the proportion of total demand allocated to the bus mode, for a given bus headway. This is presented in (13), with parameters as defined in Table 1.

$$\frac{\partial CS}{\partial h} = \frac{\gamma_w}{2}\left(\frac{e^{\frac{\alpha_3 h}{2}+M}}{e^{\frac{\alpha_3 h}{2}+M}+1}\right)\mathrm{TD(t)} = \lambda_D \frac{1}{h^2} \tag{13}$$

where

TD(t) = Total demand across all modes on route as a function of time (in passengers per minute)

h(t) = bus headway

α_3 = Logit coefficient for wait time, a measure of headway from the passenger perspective

M = Fixed component of transit mode choice representative utility

Additional insights were explored via a re-framing of (13). In (14), simplification is made by taking the negative of exponential logit coefficients with the knowledge they are negative for mode choice ($\alpha_3'$and M' are the positively signed counterparts to α_3 and M).

$$\frac{e^{-(\frac{\alpha_3' h}{2}+M')}}{e^{-(\frac{\alpha_3' h}{2}+M')}+1} = \frac{2\lambda_D}{\gamma_w \mathrm{TD(t)} h^2} \tag{14}$$

This can be simplified to the form shown in (15).

$$1 + e^{\frac{\alpha'_3 h}{2} + M'} = \frac{\gamma_w \mathrm{TD(t)} h^2}{2\lambda_D} \tag{15}$$

Equation 15 can be solved for h using numerical methods. However, a solution is not available for all possible sets of the inputs.

6 Application with Edmonton Data

With passenger arrival rate as a function of dispatch rate, one must have a context-specific function of demand that includes this variable. The logit model for Edmonton was used for the purposes of illustrating the method of calculating an optimal headway based on (13). In the Newell Policy, the passenger arrival rate, as a function of time, is used to determine headway. In the present policy formulation, the total demand on the route (including all modes) is used and the logit model applied to perform mode assignment. Total demand is represented as TD and variables in the measurable condition function – excepting total wait time – are represented by M. The Edmonton logit model represents time in units of minutes, but the present policy formulation is based on an hourly rate, thus a unit conversion was necessary to reconcile the function components. A single time period will be examined; as such, an average headway of $\bar{h}$ and average passenger arrival rate of $\bar{p}$ will be employed in the present application. This follows the derivation by Wirasinghe (1990). It will be possible to derive a time variable passenger arrival function from the Edmonton activity-based model under development, but these data are not presently available. Average wait time can be converted to dispatch rate by noting the following relationship

$$Average\ Wait\ Time = \frac{\bar{h}}{2} = \frac{1}{2\bar{g}} \tag{16}$$

Total cost functions were plotted in Fig. 5 and 6 under totally inelastic and relatively inelastic conditions, respectively, with respect to dispatch rate according to the traditional Newell method by which a minimum cost search is completed. This provides a baseline comparison between the original formulation and the modifications introduced by assuming a relatively inelastic logit passenger demand function. In the case of totally inelastic demand, it can be shown that the change in total cost - with variation of the bus headway - is equivalent to the change in consumer surplus. This is not the case with a logit demand function and it will be shown that the total cost function does not exhibit an absolute minimum. The change in consumer surplus will subsequently be derived according to (13) to obtain an optimum headway. Assumed values, for application to the Edmonton mode choice model, are summarized in Table 1.

By taking the ratio of α_5 (transit fare) to α_3 (transit wait time), one can derive a wait time cost of \$0.17/user-min or \$10.45/user-hour. Cordon count data from 2012 was used for the period 7:30 a.m. to 8:30 a.m. at the location: 104th Ave. NW east of 111th St. A total hourly auto and bus volume of 2784 was recorded at this location, with a transit bus volume of 595 passengers. Buses were observed to be 30% full across the cordon during the AM peak. There are four routes (#2, #109, #111, and #112) operating during the AM peak. The routes have headways as follows: #2 – 15 minutes, #109 – 30 minutes, #111 – 30 minutes, and #112 – 30 minutes (Edmonton Transit Service 2015). Cordon counts recorded 18 buses during the sample period, which gives an effective headway on the link of 3.3 minutes. Route #2 was used for analysis and a synthetic link produced representing the total demand for this single route. This synthetic link has an hourly bus volume of 135 passengers per hour (2.25 passengers per minute) and a total hourly demand of 620 persons per hour (10.3 persons per minute), including auto trips. It is assumed that all auto users could make the same trip by bus.

6.1 Totally Inelastic Case

Taking the assumption of passenger arrival rate being insensitive to variation in bus dispatch rate

$$\bar{z} = \frac{\gamma_W \bar{p}}{2\bar{g}} + \lambda_D \bar{g} \tag{17}$$

The optimal headway can then be expressed as

$$h^*(t) = \min \begin{cases} \left(\frac{2\lambda_D}{\gamma_W \bar{p}}\right)^{0.5} \\ \frac{c}{\bar{p}} \end{cases} \tag{18}$$

Utilizing (18), the optimal headway was determined to be 20.2 minutes (3 bus/hr) for a total demand of 135 bus passengers. The total cost function has a single minimum and no absolute maximums.

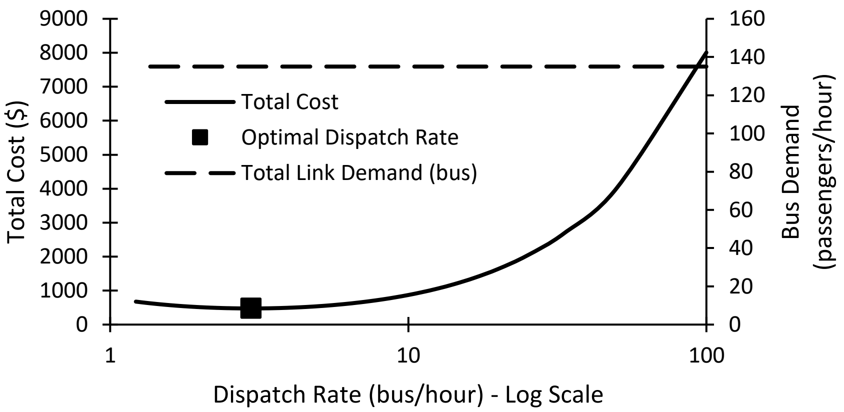

Fig. 5. Total cost for variable average dispatch rate (totally inelastic case) per 620 link users.

6.2 Relatively Inelastic Case

The demand curve increases with the dispatch rate at a decreasing rate. This approximates a constant demand at high dispatch rates; the relationship between demand and total cost follows a trend similar to that observed under totally inelastic assumptions (Fig. 6). However, at dispatch rates of fewer than 15 bus/hour, the demand rapidly dissipates. This reduces the contribution of wait time to total cost, whereas under inelastic conditions this contribution continues to increase as a constant passenger demand is subjected to increasing wait times.

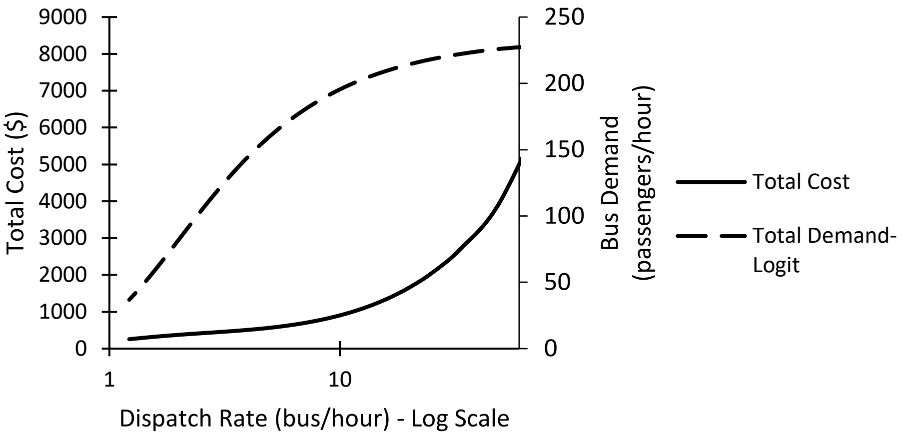

Fig. 6. Total cost for variable average dispatch rate (relatively inelastic case) per 620 link users.

Fig. 6. was developed by varying the bus dispatch rate and calculating the total cost function and bus passenger demand (for the corresponding average wait time) across a range of potential rates of dispatch. A log scale was used for dispatch rate to better illustrate the shape of the cost function. It shows a reduction in the total cost at low dispatch rates whereas, under the assumptions of Newell, the total cost increases in this regime as a fixed passenger demand faces increasing wait time costs. Lower rates of dispatch will decrease operating costs and this reduction begins to outstrip the rate of increase in wait time cost for passengers.

6.1.1 Captive Demand in the Bus Demand Function

This assumes that, when faced with increasing wait times, passengers will continue to shift to other modes or activities rather than take the bus. This raises an important issue with both the elastic and inelastic formulations. Taking demand to be inelastic to headway assumes that all users are captive passengers and will not choose another form of transportation, or other activity, when faced with long headways. Conversely, assuming elasticity for all passengers makes the assumption that all users are choice users. Neither assumption is ideal and the optimization should be based on a certain portion of users being captive and travelling by bus, regardless of the wait time cost. A "captive user" trip proportion was estimated using a combination of Edmonton and Calgary household travel survey data (ISL Consulting Inc and Banister Research & Consulting Ltd 2006; City of Calgary 2013). This analysis produced a captive share of 35% above which demand should be maintained. This is a simplifying assumption and other factors may contribute to passenger captivity such as age or other impediments.

7 Application of Consumer Surplus to Optimization Problem

In the case of totally inelastic demand, the total cost function is equivalent to the measure of consumer surplus. Each member of the fixed passenger total experiences a constant change in consumer surplus for a unit change in bus headway, which equals the negative of the additional wait time cost. When the same analysis is applied to the case of relatively inelastic demand, the number of passengers changes, such that the change in total cost does not equal the change in consumer surplus. Put differently, each passenger experiences the same wait time cost under an assumption of relative inelasticity, but the willingness to accept a higher wait time cost varies between passengers such that their consumer surplus is not uniform. Therefore, (13) should be utilized to develop a graphical solution to the optimization problem, with respect to consumer surplus rather than total cost. This can be viewed as an application of the principles outlined in Fig. 4. The results of this comparison are summarized in Fig. 9, which shows an optimal solution of 3 bus/hr, or a headway of 20 minutes. This is within the range of the 15 minute headway currently employed for route #2 in Edmonton, suggesting the current route is providing better than required service according to the inputs specified in this model.

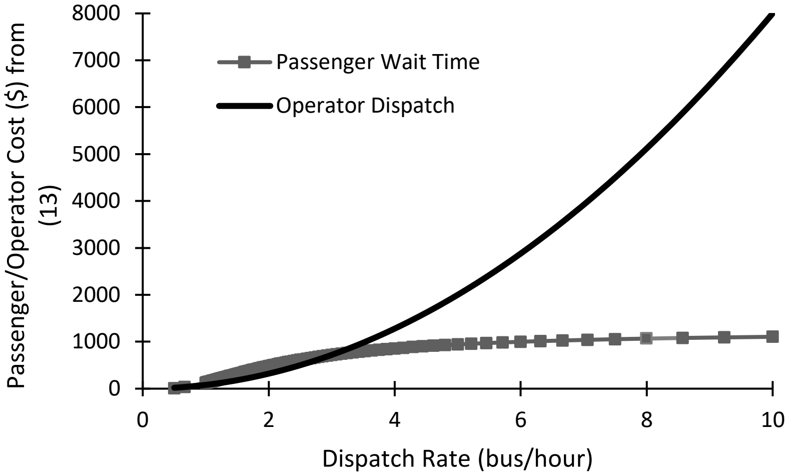

Fig. 7. Graphical solution to (13) using a relatively inelastic logit demand representation

7.1 Sensitivity to Wait Time Cost and Total Demand

It was determined, for input parameters derived for the City of Edmonton, an optimal headway exists to balance operator cost with consumer surplus from passenger wait time savings, as defined in (13). However, varying the unit wait time cost – value of time - will change this result as the rate at which people move away from bus use varies with their value of time and the quality of service. In some cases, no analytical solution exists and the solution is determined graphically using the lowest headway corresponding to captive demand of 35% of existing bus passengers. This represents a minimum derived from assumptions about captive demand, rather than an analytical solution. Public transport is deemed a public good and these results suggest the passenger requires a higher value in the analysis than that suggested by the utilities provided by the City of Edmonton.

The sensitivity of the optimization function was tested for various values of time and a higher passenger demand. The aim being to determine the range in which a minimum emerges. It was determined that, for the given route demand pool of 620 potential passengers, no local minimum occurs for values of time above about $15.00. The solution is consistently to match captive demand (dashed lines in Fig. 8 and Fig. 9) for wait time costs ranging from $25/pass-hr to $100/pass-hr.

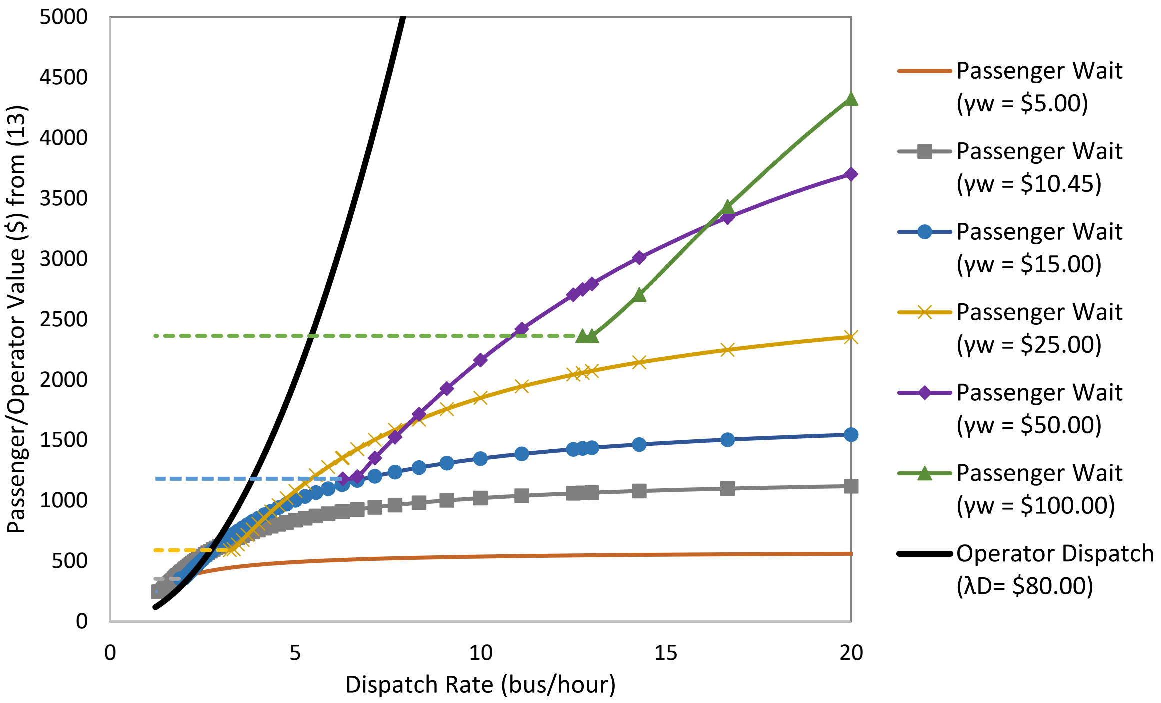

Fig. 8. Graphical solution to (13) with various values of wait time (relatively inelastic case) per potential 620 route users

For a volume of 620 users per hour and vehicle capacity of 75 passengers per bus (TRB 2014), the Edmonton-derived 620 potential route users could be fully accommodated by buses operating at headways of 7 minutes (8 bus/hr). Increasing the demand for the route should introduce a need for additional buses without requiring large proportions of total demand to travel by bus. The total demand on the route was increased by an order of magnitude (Fig. 9) to 6200 users per hour. This produced sufficient demand for bus service to exhibit optimal solutions, independent of the captive demand threshold, for all values of wait time. The shape of the function remains the same, but the total passenger cost is shifted up by the increased total route users.

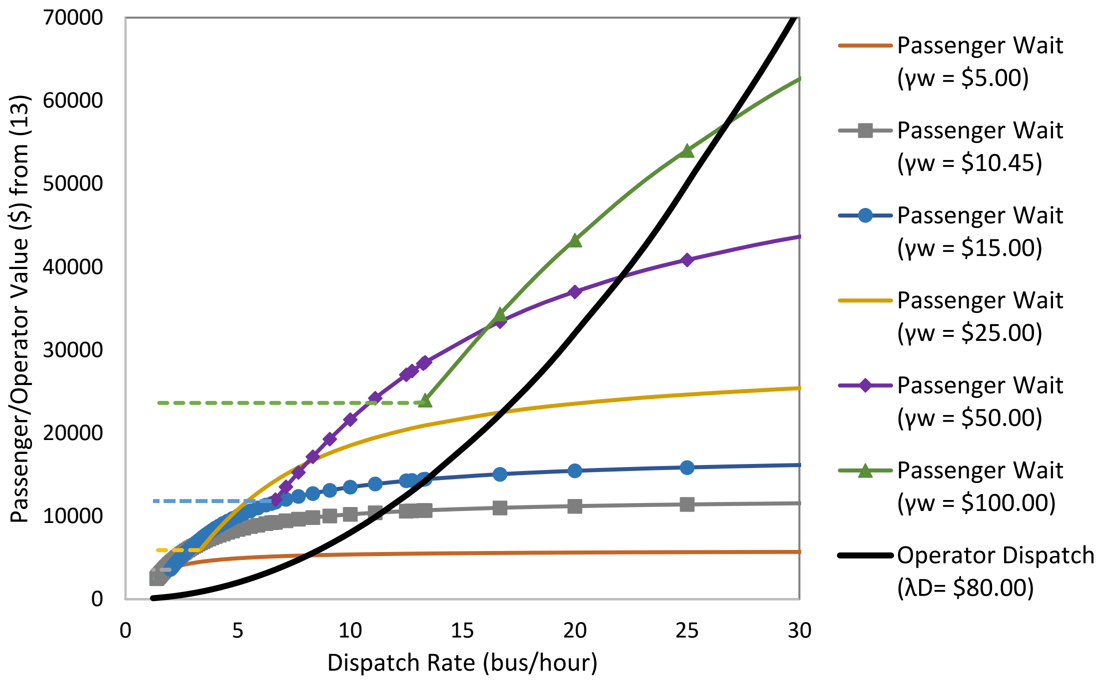

Fig. 9. Graphical solution to (13) with various values of wait time (relatively inelastic case) per potential 6200 route users

8 Conclusions

This paper aimed to develop an optimal bus dispatch policy, based on the policy developed by Newell in the 1970s, assuming elasticity of passenger demand with headway. It was proven that this model is consistent with Newell through representation of wait cost as a measure of consumer surplus. This policy was developed by maximizing the total consumer surplus – as measured by wait time cost savings – constrained by bus dispatch cost. This formulation was developed with the use of a logit mode choice utility function Several methods of system-wide optimization exist that utilize a modal utility function, but no simplified formulation exists for application to a single route. A closed form solution for headway could not be derived using a logit-based demand function. This was also the finding of past studies (Kocur and Hendrickson 1982; Ceder 1984). A headway optimization function was developed and Edmonton utility parameters applied for comparison with the totally headway-inelastic function developed by Newell. This analysis illuminated the existence of choice users, not considered in the Newell formulation. Further investigation determined that an accurate algorithm must consider passenger demand to be only relatively inelastic with bus headway, but also include a provision for captive users who do not have an alternative means of transportation. Results suggest that by using realistic values of wait time cost and demand for the Canadian context, an optimal headway cannot be determined based on optimization of competing costs. Rather, the cost of dispatching outweighs the value of time of passengers and buses should be dispatched to meet the needs of captive users. This does not fit with the "societal good" role of public transit. By considering the consumer surplus, rather than wait time cost experienced at the stop, an optimal solution can be obtained graphically. The solution developed in this research built upon the work of Newell for a single route. A simple representation of the competition between passenger and operator costs was developed, which could be applied to any standard demand function. The model was validated through a case study in Edmonton, Alberta.

The empirical test conducted could be improved via collection of a variety of additional data. Specification of an individual bus route for which there are no, or minimal, common links with other bus routes would be ideal and ensure an isolated system. This paper considered a binary choice set of travelling by bus or choosing to engage in another activity, without consideration for the potential increase in congestion resulting from bus passengers shifting to travel by private automobile. Future work could examine this function in the context of a broader definition of consumer surplus, including changes in auto user consumer surplus and road costs associated with higher auto volumes in the case of transit passengers shifting to auto at low rates of bus dispatch.

References

Balcombe, R., R. Mackett, N. Paulley, J. Preston, J. Shires, H. Titheridge, M. Wardman, and P. White. 2004. *The Demand for Public Transport: A Practical Guide*. http://www.demandforpublictransport.co.uk/TRL593.pdf.

Calgary Transit. 2002. *A Review of Bus Rapid Transit*. http://www.calgarytransit.com/pdf/brt%7B_%7Dreport.pdf.

Ceder, Avishai. 1984. "Bus Frequency Determination Using Passenger Count Data." *Transportation Research Part A: General* 18 (5–6): 439–453. doi:10.1016/0191-2607(84)90019-0.

Chang, S. K., and P. M. Schonfeld. 1993. "Welfare Maximization with Financial Constraints for Bus Transit Systems." *Transportation Research Record* 1395: 48–57.

City of Calgary. 2013. *Changing Travel Behaviour in the Calgary Region,* Transportation Planning. http://www.calgary.ca/Transportation/TP/Documents/forecasting/Changing Travel Behaviour in the Calgary Region_v1_forWeb_2013-06-04.pdf.

City of Edmonton. 2004. *Segment 16: Adult Worker Home to Work Trips,* Transportation Planning.

City of Edmonton. 2014. "Fares.", Transportation Planning. http://www.edmonton.ca/transportation/ets/fares-ets.aspx.

Edmonton Transit Service. 2015. "Route Brochures & System Maps." http://www.edmonton.ca/transportation/ets/alerts_routes_schedules_maps/route-brochures-system-maps.aspx.

Florian, Michael, R Chapleau, S Nguyen, C Achim, L James-Lefebvre, S Galarneau, J Lefebvre, and C Fisk. 1979. "Validation and Application of an Equilibrium-Based Two-Mode Urban Transportation Planning Method (EMME)." *Transportation Research Record*. http://trid.trb.org/view.aspx?id=153286.

Furth, P G, and N H M Wilson. 1981. "Setting Frequencies on Bus Routes: Theory and Practice." *Transportation Research Record: Journal of the Transportation Research Board* 818: 1–7.

Hossain, Mohammad Shafayat, John Douglas Hunt, and S. C. Wirasinghe. 2015. "Nature of Influence of out-of-Vehicle Time-Related Attributes on Transit Attractiveness: A Random Parameters Logit Model Analysis." *Journal of Advanced Transportation* 49 (5): 648–662. doi:10.1002/atr.1297.

ISL Consulting Inc, and Banister Research & Consulting Ltd. 2006. *2005 Edmonton Household Travel Survey*. Edmonton.

Kocur, G, and C Hendrickson. 1982. "Design of Local Bus Service with Demand Equilibration." *Transportation Science* 16 (2): 149–170. doi:10.1287/trsc.16.2.149.

Lam, W, and J Morrall. 1981. "Bus Passenger Walking Distances and Waiting Times: A Summer-Winter Comparison." In *RTAC Annual Conference Preprints*, 39–60. Transportation Association of Canada.

Mohring, Herbert. 1972. “Optimization and Scale Economies in Urban Bus Transportation.” *The American Economic Review* 62 (4): 591–604. doi:10.2307/1806101.

Newell, G F. 1971. “Dispatching Policies for a Transportation Route.” *Transportation Science* 5 (1): 91–105. doi:10.1287/trsc.5.1.91.

Statistics Canada. 2011. *Commuting to Work , 2011. National Household Survey,* Office for National Statistics.

TRB. 2014. *Transit Capacity and Quality of Service Manual*. http://onlinepubs.trb.org/onlinepubs/tcrp/tcrp100/part 4.pdf.

Verbas, İ Ömer, and Hani S Mahmassani. 2015. “Exploring Trade-Offs in Frequency Allocation in a Transit Network Using Bus Route Patterns: Methodology and Application to Large-Scale Urban Systems.” *Transportation Research Part B: Methodological* In Press. http://www.sciencedirect.com/science/article/pii/S0191261515001472.

Vuchic, Vukan R. 2007. “Influence of Stop/Station Spacing.” In *Urban Transit Systems and Technology*. Hoboken, N.J.: John Wiley & Sons.

Wirasinghe, SC. 1990. “Re-Examination of Newell’s Dispatching Policy and Extension to a Public Route with Many to Many Time-Varying Demand.” In *The Eleventh International Symposium on Transportation and Traffic Theory*, edited by M Koshni, 363–378. Elsevier.

Wirasinghe, SC, and G Liu. 1995. “Determination of the Number and Locations of Time Points in Transit Schedule Design - Case of a Single Run.” *Annals of Operations Research* 60 (1): 161–191. doi:10.1007/BF02031945.