The Corridor Problem

Richard Arnott*
Department of Economics
Boston College
Chestnut Hill, MA 02467
richard.arnott@bc.edu

December 2001

Abstract

Consider a corridor which connects a continuum of residential locations to the CBD (central business district) and which is subject to flow congestion. All (identical) individuals travel along the corridor from home to work in the morning rush hour and have the same work start time. Each individual decides when to depart from home so as to minimize the sum of travel time costs, time early costs, and toll costs (when applicable). This paper investigates the pattern of traffic flow over the morning rush hour and the social optimum, and considers the implications for land use and road cost-benefit analysis.

JEL code: R1, R4

Keywords: Traffic congestion, land use, monocentric city, shadow rent

*The author would like to thank seminar participants at U.B.C., Dong Hua University, Kyoto University, the University of Tokyo, and the Free University of Amsterdam, especially Yasushi Asami, for helpful comments, and Junling Wang, Xiaomeng Yang, and Lada Lobas for assistance in programming the solution algorithm.

The Corridor Problem

1 Introduction

Consider a single road which connects a continuum of residential locations to the CBD (central business district) and which is characterized by flow congestion. Population density and road width are exogenous and vary along the road. All (identical) individuals travel along the road from home to work in the CBD in the morning rush hour, and have a common work start time. Late arrival is not permitted. Consequently, each individual decides when to leave home so as to minimize the sum of travel time costs, time early costs, and toll costs (where applicable).

This paper investigates the pattern of traffic flow along the corridor over the morning rush hour in both the untolled equilibrium and the social optimum. It then considers the implications of the results for equilibrium residential land use and the cost-benefit analysis of road capacity.

This "corridor problem" is a central problem in the theoretical literature on traffic congestion and land use, tieing together three previously unconnected branches of the literature. The first is the 70's literature on traffic congestion in the monocentric city model (Solow and Vickrey (1971), Dixit (____), Solow (1972,1973), Kanemoto (1976), and Arnott (1979)). The last four papers considered the optimal allocation of land to a radial congestible road (circumferential travel being costless) in the basic monocentric city model, in both the first best where optimal congestion tolls are applied and in the second best where they are not. The models were static, and assumed that an individual's travel cost between distances x and $x + dx$ along the road away from the CBD depends on the "volume-capacity" ratio there, with volume treated as the number of individuals living beyond x. All the aforementioned contributors to the literature recognized this as an uncomfortable, "reduced-form" assumption — a static approximation of an intrinsically dynamic phenomenon. Providing an explicit dynamic treatment would have required both treating individuals' trip-timing decisions and tackling the mathematics of non-stationary traffic flow theory, which seemed intractable.

Vickrey made two breakthroughs in his 1969 paper. He recognized that an individual decides when to travel so as to minimize trip price, where trip price includes not only the money and time costs of the trip, as well as tolls payable, but also the cost of arriving at the destination inconveniently early or late — what has come to be known as "schedule delay cost". He also recognized that tractability could be achieved by modelling congestion as queues behind bottlenecks. His model in that paper, the first tractable model of the dynamics of rush-hour congestion, has come to be referred to as the "bottleneck model".

The second branch of the literature to which the corridor problem is related is the elaboration of the bottleneck model. The bottleneck model is aspatial. Arnott, de Palma, and Lindsey(1993) took a first step towards making the bottleneck model spatial by examining two bottlenecks in series with two entry points and a single exit point. Arnott and Lindsey (in rough notes) attempted to generalize the analysis to an arbitrary number of bottlenecks in series, with entry between each pair of bottlenecks. Unfortunately, the solution entailed cases, with the number of cases increasing rapidly in the number of bottlenecks.

The third branch of the literature to which the corridor problem is related are papers by Yinger (1993) and Ross and Yinger (____). Yinger (1993) explored the traffic pattern along a corridor under the *assumption* that all individuals at a particular location depart for work at the same time. Ross and Yinger (____) demonstrated that neither this assumption nor the assumption that at each location individuals depart at a constant rate over a rush hour of fixed duration (which they refer to as the Solow assumption) is consistent with the "trip-timing condition" that no individual can reduce her trip cost by changing departure time, but they stopped short of deriving the temporal-spatial pattern of departures consistent with the trip-timing condition.

Finally the coridor problem extends Newell (1988),which examined the equilibrium and optimum temporal-spatial patterns of trafic congestion on a point-input, point-output section f road during the morning rush hour, with flow congestion.

Since the first branch of the literature provides a rich treatment of the interaction between land use and transport congestion, which is of considerable theoretical and practical interest, this paper's principal contribution is to investigate how good an approximation the static treatment of congestion in the 70's literature is to a proper treatment which incorporates the physics of traffic flow and accommodates the trip-timing condition, and whether the qualitative results of the 70's literature are robust to a proper treatment of congestion. Unfortunately, the problem of analytical intractability remains. The paper does succeed in providing a qualitative characterization of the equilibrium departure pattern, but going beyond this resorts to numerical solution.

Section II gives a precise statement of the corridor problem with fixed land use, and provides a qualitative characterization of the equilibrium and optimal departure patterns for the general case. Section III describes an intuitive algorithm that generates approximate numerical solutions to the corridor problem, and applies the algorithm to solve for trip costs as a function of distance from the CBD for the untolled equilibrium and a pseudo-social optimum, and compares the results to those that would be obtained with the static treatment of congestion. It then performs a similar exercise with respect to the shadow rent on land in road use [subsequent sections which are not yet completed will endogenize the land use pattern, as was done in the seventies literature].

2 The Corridor Problem with Fixed Land Use: Flow Congestion

2.1 Description of model

A notational problem is immediately encountered on how to index locations. In the monocentric city model, locations were naturally indexed according to distance from the CBD. However, the bottleneck and related models which incorporate trip timing have focused on the *morning* commute [1], in which travel is from residential locations to the CBD, and it is natural to describe cars as going "forward" (in the direction of an increasing locational index) rather than backwards. The choice has been made to index locations in terms of distance from the outer boundary of residential settlement. This presents no difficulty in the next three sections where the land use pattern is fixed, but will subsequently, in those sections where land use is endogenous.

Documents/Nika/Arnott/figures/Slide3.wmf

Consider a traffic corridor which connects a continuum of residential locations, "the suburbs", to the CBD, which lies at the eastern end of the corridor. Location is indexed by x, distance from the outer boundary of residential settlement to the CBD, which is located at $\bar{x}$; $w(x)$ denotes the exogenous width of the road at x ; and $N(x)\,dx$ the exogenous number of individuals living between x and $x + dx$.

Each morning all individuals drive from their homes to work in the CBD, which starts at the common time t^*. Late arrivals are not permitted, and the common travel cost function is[2]

$$C = \alpha(traveltime) + \beta\,(timeearly)\,,$$

where α is the value or cost of travel time, and β the value or cost of time early at work.

[1] In the evening commute, it is not clear whether schedule delay costs should be defined with respect to a desired departure time from work or a desired arrival time at home. See de Palma and Lindsey [____] for the state of the art with respect to the evening-commute bottleneck model.

[2] The analysis can be extended straightforwardly to treat more general travel cost functions. This one is employed because it greatly simplifies the analyisis and facilitates exposition.

Let t denote time, and $T(x,t)$ the travel time of an individual who departs from x at time t. Then

$$C(x,t) = \alpha T(x,t) + \beta \left(t^* - (t + T(x,t))\right), 1 \qquad (1)$$

where $t + T(x,t)$ is the individual's arrival time at work and hence $t^* - (t + T(x,t))$ time early.

The final feature of the economy to be described is the technology of traffic congestion. Classical flow congestion is assumed, which combines the equation of continuity with an assumed technological relationship[3] between density and velocity and appropriate boundary conditions. The equation of continuity is simply a statement of conservation of mass, that the change in the number of cars on a section of road of infinitesimal length equal the inflow minus the outflow. Letting $D(x,t)$ be the density of cars (per unit area) at location x at time t, and $v(x,t)$ the corresponding velocity, the equation of continuity is

$$D_t w = -v_x w D - v w' D - v w D_x + n, 2 \qquad (2)$$

where subscripts denote partial derivatives, $w' \equiv \frac{d(w(x))}{dx}$, and $n(x,t)$ is the entry rate onto the road at location x at time t. The relationship between velocity and density can be written as

$$v(x,t) = V(D(x,t)), 3 \qquad (3)$$

with $V' \equiv \frac{dV(D)}{dD} < 0$. Substituting (3) into (2) yields

$$D_t w = -V' D_x w D - V w' D - V w D_x + n, 4 \qquad (4)$$

which is a partial differential equation for $D(x,t)$. The appropriate boundary conditions are simply that the road be empty prior to the start of the rush hour.

[3]From the Fundamental Identity of traffic flow theory that flow$\equiv$ density x velocity, this technological relationship can be expressed as a relationship between any pair of flow, density, and velocity. The density in the fundamental identity is per unit length and so corresponds to Dw.

2.2 Characterization of the no-toll equilibrium

There are two equilibrium conditions. The first is that everyone commute, the second the trip-timing condition that no individual can reduce her trip cost by departing at a different time.

Let $\mathcal{D}(x,t)$ denote the set of (x,t) for which departures occur in equilibrium. The everyone-commutes condition is that

$$\int_{\left(x',t\right)\in\mathcal{D}(x',t)} n\left(x',t\right)dt = N\left(x'\right) \quad forall x' \in (0,\bar{x})\,5 \tag{5}$$

at each location, the integral of the departure rate over the set of departure times for that location equals the population at that location. Let $p\left(x\right)$ be the equilibrium trip price at location. The trip-timing condition is then

$$C\left(x,t\right)\{.= p\left(x\right) \qquad for\left(x,t\right)\in\mathcal{D}(x,t) \geq p\left(x\right) \qquad for\left(x,t\right)\notin\mathcal{D}(x,t),6 \tag{6}$$

which states that at no location can the trip price be reduced by travelling outside the departure set at that location.

A *no-toll equilibrium* is then a departure pattern $n\left(x,t\right)\geq 0$ and a trip price function $p\left(x\right)$ such that:

$i)$ (5) is satisfied

ii) $\left.\begin{array}{c} n\left(x,t\right)\left(C\left(x,t\right)-p\left(x\right)\right)=0 \\ C\left(x,t\right)\geq p\left(x\right)\end{array}\right\}$ for all $\left(x,t\right)$

and $C\left(x,t\right)$ is given by (1), with $T\left(x,t\right)$ obtained from (4) with the boundary conditions that there be no traffic on the road either before the start of the rush or after t^*. Condition *ii)* is written in variational inequality form[4].

[4]Rather than impose the condition that there be no traffic on the road after t^*, one can define equilibrium as the limiting equilibrium as $\gamma\uparrow\infty$, with the cost function

$$C\left(x,t\right)=\alpha T\left(x,t\right)+\max\left(\beta\left(t^*-\left(t+T\left(x,t\right)\right)\right),\gamma\left(t+T\left(x,t\right)-t^*\right)\right)$$

where γ is the cost of time late; late arrival is infinitely expensive rather than simply disallowed.

Throughout the rest of the paper it is assumed that $\alpha > \beta$, which is supported by empirical observation (Small (1982)). This assumption ensures that $n(x,t)$ has no mass points, and that T, T_t, T_x, and T_{xt} are continuous functions.

From (1) and (6),

$$p(x) = \alpha T(x,t) + \beta(t^* - t - T(x,t)) \quad for\,(x,t) \in \mathcal{D}(x,t), 7a \tag{7}$$

implying that

$$T(x,t) = \frac{p(x) - \beta(t^* - t)}{\alpha - \beta} \qquad for\,(x,t) \in \mathcal{D}(x,t).7b \tag{8}$$

Thus, at each location, over the departure set at that location travel time increases linearly in departure time at the rate $\frac{\beta}{\alpha-\beta}$.

It will prove convenient at this point to make the transformation of variables

$$a(x,t) = t + T(x,t), 8 \tag{9}$$

where $a(x,t)$ is the arrival time at the CBD of an individual who departs location x at time t. The inverse transformation is

$$t(x,a) = a - \hat{T}(x,a), 8$$

$'$ (10)

which relates departure time to arrival time. The trip-timing condition, expressed in terms of arrival time, is

$$p(x) = \alpha\hat{T}(x,a) + \beta(t^* - a) \qquad for\,(x,a) \in \mathcal{A}(x,a), 9 \tag{11}$$

where $\mathcal{A}(x,a)$ is the set of (x,a) for which the arrival rate is positive. The advantage of working in terms of arrival time is that $\hat{T}\left(x,a'\right)$ tracks the cohort of vehicles that arrive at time a'. Then

$$\hat{T}(x+dx,a) = \hat{T}(x,a) + \frac{dx}{v\left(x, a - \hat{T}(x,a)\right)}, 10a \tag{12}$$

since a vehicle which arrives at a passes location x at $a-\hat{T}\left(x,a\right)$. Then

$$\begin{aligned}\hat{T}_x\left(x,a\right) &= \frac{1}{v\left(x,a-\hat{T}\left(x,a\right)\right)} \\ &= \frac{1}{V\left(D\left(x,a-\hat{T}\left(x,a\right)\right)\right)}.10b\end{aligned} \tag{13}$$

Differentiation of (9) with respect to a and x yields

$$.\}\hat{T}_a = \frac{\beta}{\alpha}\hat{T}_{ax} = 0 for\left(x,a\right)\in int\mathcal{A}\left(x,a\right), 11a, b \tag{14}$$

while differentiation of (10b) with respect to a yield

$$\hat{T}_{xa}\left(x,a\right) = -\frac{V^{\prime}D_t\left(1-\hat{T}_a\right)}{V^2}12 \tag{15}$$

Since V and $V^{\prime}$ are strictly positive, (11) and (12) together imply that

$$D_t\left(x,a-\hat{T}\left(x,a\right)\right) = 0 \quad for\left(x,a\right)\in int\mathcal{A}\left(x,a\right)13a \tag{16}$$

and

$$D_t\left(x,t\right) = 0 \quad for\left(x,t\right)\in int\mathcal{D}\left(x,t\right).13b \tag{17}$$

Eq. (13b) states that traffic density is constant at a particular location over the interior of each connected subset of the departure set at that location, which implies from (4) that the departure rate too is constant over the interior of each connected subset of the departure set at that location.

Documents/Nika/Arnott/figures/Slide2.wmf

A departure pattern consistent with these results is displayed in Figure 2. The function $t_0\left(x\right)$ indicates the time of the first departure from x, and the function $t_{\ell}\left(x\right)$ the time of the last departure. At x, departure occurs at the rate $n(x,t) = \frac{N_x}{t_{\updownarrow}(x)-t_o(x)}$ from $t_o(x)$ to $t_{\updownarrow}(x)$. The departure set is the area between $t_0\left(x\right)$ and $t_{\ell}\left(x\right)$. A dashed-and-dotted lines indicates the time line of a particular cohort of cars, with slope equal to the inverse of velocity. Cohort (1) contains cars that enter the road at the rate $n\left(x,t\right) = \frac{N(x)}{t_{\ell}(x)-t_0(x)}$ from

$x = 0$ to $x = x_1$, with no cars entering closer to the CBD than x_1. Cohort (2) departs after cohort (1). The departure pattern is the same for cohort (2) as for cohort (1) up to x_1, but differs from cohort (1) in that entry occurs from x_1 to x_2 as well. The entry of these cars increases the travel time of all the cars entering the road from $x = 0$ to $x = x_1$ *by the same amount* Succeeding cohorts follow the same pattern, with travel conditions unchanged from the boundary of the city to the location closest to the CBD at which entry occurs, with this location moving continually closer to the CBD at a rate such that the travel time of all cars entering the road at more distant locations increases at the rate $\frac{\beta}{\alpha-\beta}$. The final cohort contains cars which enter from all locations and arrives at the CBD at t^*.

The departure pattern satisfies the condition that everyone commute, since at any location x departures occur at the rate $\frac{N(x)}{t_\ell(x)-t_0(x)}$ for a period of time $t_\ell(x) - t_0(x)$. Furthermore at x: i) entry earlier than $t_0(x)$ results in higher trip costs than entry between $t_0(x)$ and $t_\ell(x)$; ii) trip costs are the same for entry at any time between $t_0(x)$ and $t_\ell(x)$, the increase in travel time costs being exactly offset by the decrease in time early costs; and iii) entry later than $t_\ell(x)$ results in late arrival, which is not permitted. Thus, the indicated departure pattern satisfies the trip-timing condition too.

Such a departure pattern will be referred to as a *horn* departure pattern since its departure set (in $x - t$ space) has the shape of a horn — see Figure 2.

It can be shown that the indicated departure pattern is the unique equilibrium departure pattern by developing an exhaustive typology of alternative departure sets satisfying the everyone-commutes condition, and constructively demonstrating that each violates the trip-timing condition [5].

Documents/Nika/Arnott/figures/Slide1.wmf

Additional insight may be gained by displaying the equilibrium arrival set $\mathcal{A}(x, a)$. In x-a space, cohorts are characterized by horizontal lines since all cars in a cohort arrive

[5]Perhaps there is a theorem from the theory of partial differential equations which can be used to establish uniqueness.

at the CBD at the same time, wherever they entered the road. The functions $a_0(x)$ and $a_\ell(x)$ give the times of the first and last arrivals from x at the CBD. Since the last cohort contains cars from every location, $a_\ell(x)$ is a horizontal line at t^*. Thus, the equilibrium is fully characterized by the function $a_0(x)$.

Much ink has been spent on hypercongestion-travel on the positively-sloped portion of the flow-velocity curve. From the perspective of the corridor problem, there is nothing especially paradoxical about hypercongestion in an untolled equilibrium. Consider the extreme case where the negatively-sloped portion of the flow-velocity curve is in fact completely vertical. To satisfy the trip-timing condition, travel time must be higher for later cohorts, and this would require that travel be hypercongested for all (x, t) for which $n(x, t) > 0$. Thus, hypercongestion with flow congestion is completely analogous to queuing with bottleneck congestion. Whether hypercongestion can arise at the social optimum is considered later.

3 Characterization of the social optimum

Consider a section of road between x and $x + dx$. The cost of travel time on that section of road between t and $t + dt$ equals the flow of cars that traverse that section of road in that time interval, $f(x, t)\, dt$, times the time it takes each car to traverse that section of road times the value of travel time, $(f(x, t)\, dt)\left(\frac{dx}{v(x,t)}\right)\alpha$. From the fundamental identity of traffic flow, flow=density per unit length times velocity. Since density per unit length equals density per unit area times width, the cost of travel time on that section of road over that interval of time is simply $D(x, t)\, w(x)\, dx\ dt$. Thus, total travel time costs on the road are

$$TTC = \alpha_{\mathcal{T}(x,t)} D(x, t)\, w(x)\, dxdt, 14a \tag{18}$$

where $\mathcal{T}(x, t)$ is the set of (x, t) for which there are cars on the road. The cost of time early for arrivals at the CBD between t and $t + dt$ is $(f(\bar{x}, t)\, dt)(t^* - t)\beta$, so that total time early costs are

$$TEC = \beta \int\limits_{\mathcal{A}(0,t)} D(\bar{x},t)\, w(\bar{x})\, V(D(\bar{x},t))\, (t^* - t)\, dt.14b \tag{19}$$

The optimal control problem is therefore to choose $\langle n(x,t) \rangle$ to minimize $TTC + TEC$, subject to the everyone–commutes constraint, a non-negativity constraint on $\langle n(x,t) \rangle$, a constraint that $D(\bar{x},t) = 0$ for $t > t^*$, which ensures no late arrivals, the equation of motion for traffic flow on the road, (4), and the boundary condition that there be no traffic on the road prior to $t_0(0)$. $n(x,t)$ is the control variable, and $D(x,t)$ the state variable.

$$< D(x,t) >< n(x,t) > \min \Lambda = \alpha \int \int D(x,t)\, w(x)\, dx\, dt + \beta \int D(\bar{x},t)\, w(\bar{x})\, V(D(\bar{x},t))\, (t^* - t)\, dt$$

s.t. *i)* (5)

ii) (4)

iii) $n(x,t) \geq 0$ (15)

iv) $D(\bar{x},t) = 0 \qquad$ for $t \geq t^*$

Robson (____) employs a heuristic solution method to solve an analogous optimal control problem. First, x is held fixed, and the problem treated as a regular optimal control program with an ordinary differential equation constraint, with time as the running variable. Then t is held fixed, and the problem treated as a regular optimal control program with an ordinary differential equation constraint, with location as the running variable.

> Solution to be determined. Hypercongestion, which occurs if an increase in density decreases flow — i.e. $DV' + V < 0$- to be investigated. Conjecture: No hypercongestion occurs (though it can occur at the SO on a network — analogous to de Palma and Jehiel)

If the social optimum can be achieved through tolling, each individual should face a trip price equal to marginal social cost:

$$p(x) = \frac{\partial \Lambda}{\partial N(x)}, 16 \tag{20}$$

where $\frac{\partial \Lambda}{\partial N(x)}$ is given by the shadow price on (5). The corresponding optimal Pigouvian toll would be

$$\pi^*(x,t) = p^*(x) - \alpha T^*(x,t) - \beta(t^* - (t + T^*(x,t))).17 \tag{21}$$

4 IV. The Corridor Problem with Fixed Land Use: "Buses"

4.1 Model Specification

Unfortunately the corridor problem with flow congestion does not admit closed-form solutions. Solving the corridor problem for particular examples requires the use of the computer, which in turn requires discretizing the problem. Discretization entails, among other things, converting a partial differential equation into a partial difference equation. There are many ways of discretizing a partial differential equation, and different methods may give rise to solutions with qualitatively different properties. In the rest of the paper, a particular discretization is employed that was chosen on the basis of its intuitiveness.

It is assumed that "buses" depart at regular intervals from the boundary of residential settlement and travel to the CBD, and that individuals board at regularly-spaced bus stops, indexed by $i = 1, ..., I$, from the boundary to the CBD. The congestion technology of the buses is more akin to flow congestion than actual bus congestion, however. Loading and unloading passengers takes no time, and the speed of a bus depends on the ratio of its number of passengers to its width or capacity, which is constant between one bus stop and the next but varies across stops. It is furthermore assumed that the technology of congestion takes the Vickrey form; in particular the time it takes to travel from bus stop i to bus stop $i+1$ on bus j is given by

$$\tau_j^i = c_0 + c_1 \left(\frac{P_j^i}{w^i} \right)^\gamma,$$

where c_0, c_1 and γ are constants, P_j^i is the number of passengers travelling on bus j from stop i to stop $i+1$, and w^i is the width of the road from stop i to stop $i+1$. Employing the Vickrey congestion function has several advantages: it is familiar and its properties well-understood; it leads to simple solution of a pseudo social optimum; and traffic does not get stuck out of equilibrium since travel time is finite for any finite number of passengers. Its principal disadvantage is that it assumes that the elasticity of private congestion (the excess of travel time over free-flow travel time) with respect to the "volume-capacity" ratio is constant, whereas it is typically increasing in the ratio.

The travel cost function has the form employed in the previous section, (1)

4.2 No-toll equilibrium

An argument analogous to that of the previous section establishes that, except for problems introduced by the discreteness of buses and bus stops: *i)* the first bus picks up passengers at the first bus stop, and possibly succeeding bus stops, without skipping stops; *ii)* bus $j+1$ boards passengers at all the bus stops that bus j did, and possibly succeeding stops, without skipping; and *iii)* over a stop's departure set, an equal number of passengers board each bus. The last bus therefore picks up passengers at all stops.

Due to the discreteness of buses, the first bus may be only 'partly utilized' — which will be defined shortly. The solution algorithm has two parts. The first solves for the equilibrium number of buses, the second for the full equilibrium. In the first part the solution algorithm assumes the first bus to be fully utilized, and then solves for the minimum number of buses such that the last bus is only partly utilized. In particular: J denotes the number of buses. Start with $J = 1$. Board all passengers. If the trip time from $i = 1$ (the departure stop) to the depot at CBD (stop $I + 1$) is more than $\frac{\beta s}{\alpha-\beta}$ higher than that of bus zero which carries no passengers, where s is the spacing between bus departures, set $J = 2$, else $J = 1$ is the equilibrium number of buses. With $J = 2$, board the first bus *per* the above results until trip time is $\frac{\beta s}{\alpha-\beta}$ higher than that of bus zero. Put all remaining passengers on the second bus. If its trip time is less than $\frac{2\beta s}{\alpha-\beta}$ higher than that of bus zero, equilibrium entails one fully utilized, and one partly utilized bus, else set $J = 3$, - - -. The second part of the algorithm assumes that the last bus arrives at t^* and recognizes that it is the first and not the last bus that is fully utilized. In particular, it solves for θ such that trip time on the first bus is $\frac{\theta\beta s}{\alpha-\beta}$ higher than that on bus zero, on the second bus is $\frac{(1+\theta)\beta s}{\alpha-\beta}$ higher - - -, and such that bus J picks up all remaining passengers and trip time is $\frac{(J-1+\theta)\beta s}{\alpha-\beta}$ higher than that on bus zero.

Recall the way in which the algorithm works. The first part determines the equilibrium number of buses, and the second the θ associated with the first bus. The boarding pattern satisfies three equilibrium conditions: *i)* the travel time of bus j, for $j = 2, ..., J$, from the first bus stop to the depot is $\frac{\beta s}{\alpha-\beta}$ higher than that of the previous bus; *ii)* at each bus stop, the departure rate is constant over that stop's departure interval; and *iii)* the first bus picks up passengers at the bus stops furthest from the depot, the second bus passengers at those bus stops and the next bus stops furthest away from the CBD, and so on.

Now turn to Table 1, which describe the base case equilibrium. The parameters are chosen to provide sensible results for a long, narrow town nine miles long and one mile wide, with road width .2 miles at all locations, and with 10,000 households per mile. Bus stops are equally spaced one mile apart. One way interpret is that all households living between nine and eight miles from the depot board at bus stop 1, etc; with the final mile being occupied by the CBD and with the depot at the very edge of the city — perhaps where the port is located. Free-flow travel speed is 20 m.p.h and the congestion technology is such that the travel time between stops i and $i+1$ (with i.e. 9 denoting the depot) on bus j, τ_j^i, is

$$\tau_j^i = .05 + \left(.05x10^{-10}\right)\left(\frac{P_j^i}{w^i}\right),^{2\cdot 0}$$

where w^i is the width of the road between stops i and $i+1$, and P_j^i the number of passengers on the bus between stops i and stop $i+1$ on bus j. Buses depart every .1 hrs., with the first bus departing at such a time that the last bus arrives at the depot at t^*. The value of travel time is \$6.00/hr. and of time early is \$4.00/hr.

Table 1A: Bus Corridor

	1	2	3	4	5	6	7	8	
n_1^i	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$	24.11×10^2	0.00	0.00	0.00	0.00	124.11×10^2
n_2^i	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$	37.95×10^2	50.00×10^2	50.00×10^2	13.09×10^2	0.00	251.04×10^2
n_3^i	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$	37.95×10^2	50.00×10^2	50.00×10^2	86.91×10^2	100.00×10^2	424.86×10^2
p^i	5.49	5.19	4.86	4.48	4.04	3.47	2.76	1.63	
ρ^i	- 57.9×10^3	17.2×10^3	25.7×10^3	39.2×10^3	87.3×10^3	142.3×10^3	247.4×10^3	445×10^3	
$\widehat{\rho}^i$	0.8×10^3	6.6×10^3	22.3×10^3	55.5×10^3	125.4×10^3	239.3×10^3	496.2×10^3	952.2×10^3	
	T_3^1	θ	$t^0 - t_1^1$	TC	TFC	$TCTC$	TEC		
	0.916	0.579	1.116	31.91×10^4	10.08×10^4	15.84×10^4	5.99×10^4		

Table 1B: Solow

	1	2	3	4	5	6	7	8
n_1^i	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$
n_2^i	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$
n_3^i	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$	$33.\dot{3}\times10^2$
p^i	4.10	3.79	3.46	3.08	2.65	2.14	1.54	$0.8\dot{3}$
ρ^i	0.8×10^3	6.6×10^3	22.3×10^3	52.9×10^3	103.3×10^3	178.6×10^3	283.7×10^3	423.5×10^3
$\widehat{\rho}^i$	0.8×10^3	6.6×10^3	22.3×10^3	52.9×10^3	103.3×10^3	178.6×10^3	283.7×10^3	423.5×10^3
	T_3^1	θ	$t^0 - t_1^1$	TC	TFC	$TCTC$	TEC	
		n.a	0.883	21.60×10^4	11.52×10^4	10.08×10^4	n.a	

Table 1: Base case: $c_0 = 0.05$, $c_1 = 0.05\times10^{-10}$, $\gamma = 2.0$, $s = 0.1$, $\alpha = 6.0$, $\beta = 4.0$, $v_f = 20$, $\delta = 1.0$

$N^{\cdot} = 10,000 \quad 10,000 \quad 10,000 \quad 10,000 \quad 10,000 \quad 10,000 \quad 10,000$

$w^{\cdot} = \quad 0.2 \quad 0.2 \quad 0.2 \quad 0.2 \quad 0.2 \quad 0.2 \quad 0.2$

Notes: i) Units: miles, hours and dollars

ii) Costs are per rush hour. To convert to annualized costs multiply by 400 (2 rush hour trip per day, 200 work days per year)

iii) The tolerance on θ was 0.001 and double precision was used.

Table 1A depicts the no-toll equilibrium for the bus–corridor problem. There are three buses in equilibrium. and $\theta = .579$. The first bus departs from stop 1 at $t_1^1 = t^* - 1.116$, picking up 3333.3 passengers there and taking.$05 + (.05x10^{-10})\left(\frac{3333.3}{.2}\right)^{2.0} = .05138$ hors to travel to stop 2. There it picks up 3333.3 passengers and takes $.05+(.05x10^{-10})\left(\frac{6666.6}{.2}\right)^{2\cdot 0} =$ .05 hours to travel to stop 3, . . . It picks up its last passengers at stop 4, and with 12,411 passenghers on board travels from there to the depot where it arrives at $t_1^1 + T_1 = (t^* - 1.116) + \frac{\beta\theta s}{\alpha-\beta} = t^* - 1.116 + .516 = t^* - .6$. The second bus departs at $t^* - 1.016$, picks up the same number of passengers as bus 1 at stops 1,2, and 3, and more passengers at stop 4. It also picks up passengers from stops 5,6, and 7, which bus 1 did not, and travels from stop 7 to the depot with 25,104 passengers on board, arriving at the depot at $t_2^1 + T_2 = (t_1^1 + s) + \frac{\beta(1+\theta)s}{\alpha-\beta} = t^* - .3$. The final bus picks up the same number of passengers as did bus 2 at bus stops 1,...,6, picks up more passengers at stop 7, picks up all the households at stop 8, and travels from there to the depot with 42,486 passengers on board, arriving at the depot at t^*.

This departure pattern ensures that the trip price at a particular bus stop is the same on all buses for which there are departures at that stop. The trip price at bus stop 1 is \$5.49, etc. Total travel costs are are simply calculated as $TC = \sum_i p^i N^i =$ \$31.91 x 10^4.

Several other aggregates are calculated. TFC are total free-flow travel costs $-$ what travel costs would be if there were no congestion and everyone arrived at work on time. The costs due to congestion ate $TC - TFC$, which can be decomposed into total congestion travel time costs, $TCTC-$ the increase in travel time costs due to congestion$-$ and total time early costs, TEC.

Figure 4 provides a diagrammatic depiction of the equilibrium pattern of travel in x-t space. The lower bold line is the locus of earliest departure times at the various bus stops, the upper bold line the corresponding locus of latest departure times (which since the last bus boards passengers from all stops, coincides with the time line of the last bus). The dashed lines are the time lines for the various buses. The slope of a bus' time line at a particular (x, t) gives the reciprocal of its velocity when is passes location x at time t. Buses depart at intervals of .1 hr and arrive at the depot at intervals of .3 hrs.

Documents/Nika/Arnott/figures/Slide4.wmf

The $\{\rho^i\}$ give the shadow rents on land in road use per rush hour between stops i and $i+1$. ρ^i was calculated by widening the road from .200 to .201 between stops i and $i+1$, with the road width being held fixed at .200 at all other locations, calculating the corresponding equilibrium, computing the reduction in total costs from the road widening, and then multiplying by 1000 to give the rent in \$/ml.^{2}. The $\{\hat{\rho}^i\}$ gives the corresponding shadow rent, *holding the boarding pattern as it was in the base case* , but postponing the departure time of the first bus such that the last bus continues to arrive on time,[6] which is here termed the pseudo shadow rent. Four points are of particular interest:

i) Both the shadow rent and pseudo shadow rent fall off with distance from the CBD, as simple intuition would suggest since for every bus the road is at least as congested at more central locations.

*ii)*The pseudo shadow rents are at least as high as the actual shadow rents reflecting that some of the benefits of the road expansion are dissipated due to the change in the boarding pattern included by the road expansion. This phenomenon is typical of second-best economies. Arnott (1979) termed the benefits from the road expansion, not allowing general equilibrium adjustments to it, the *direct benefits*, and the loss in benefits due to the general equilibrium adjustments, the *indirect costs*. Here there is only one general equilibrium adjustment — the change in the boarding pattern; in particular, it becomes even more excessively concentrated relative to the social optimum (which shall be examined subsequently).

iii) The indirect costs from a road expansion may exceed the direct benefits, leading to a negative shadow rent; in the example, this occurs only at the bus stop 1.

iv) The shadow rent on land in road use at bus stop 8 is very high, \$695 per acre-rush-hour, which with 2 rush hours per day, 200 working days a year, and a discount rate of 5%,

[6]Since the real widening decreases the travel time of the last bus more than that of earlier buses, total time early costs are reduced, as well as total congestion travel time costs.

translates into a shadow value of land in road use of \$5.56 x 10^6 *per* acre; which suggests that in a fuller model which endogenizes residential location[7] it will be (second-best) optimal to allocate almost all and close to the CBD to roads.

Finally, note that the assumption that everyone starts work at the same time tends to inflate total travel costs. In response to congestion, some firms change their work start times, while others offer flexible working hours, which would reduce congestion. It would be interesting to extend the model à la Henderson (____) to endogenize work start times. One could then investigate whether households at locations further distant from the CBD choose to start work earlier or later; the natural conjecture is earlier since they have further to travel under congested conditions, and hence should derive greater benefit from early departure. Another useful extension would be to allow late arrival. The extension to treat congestion with a realisitc pattern of employment locations in a two-dimensional city will be neccessary before the model can be confronted with real world data, but will be difficult.

Table 1B displays the equilibrium for the same set of parameters but for the Solow model. The Solow model ignores time early costs, and assumes that at each location traffic flow is constant over a rush hour of fixed duration. It is not clear exactly what this corresponds to in terms of the bus/corridor model. Here it is treated as corresponding to the bus/corridor model with three buses, on each of which a third of the households at each bus stop are boarded. Trip price is lower in this interpretation of the Solow model than in the bus-corridor model for two reasons: first, the Solow model ignores time early costs, and second the departure pattern (conditional on the number of buses) is efficient — due to the convexity of travel time function, efficiency entails the same boarding pattern for each bus. Total congestion travel time costs are correspondingly lower. In the Solow model, since the departure pattern is exogenous, the indirect costs associated with widening the road are zero, and consequently the shadow rents on land in road use and the pseudo shadow rents coincide. Because its departure pattern is efficient, the Solow equilibrium entails less

[7]Such a model will be developed in subsequent papers

congestion than the bus/corridor equilibrium. As a result, the reduction in congestion costs from widening the road — holding fixed the respective departure patterns — are lower in the Solow than in the bus/corridor equilibrium, but still everywhere positive. However, because widening the road entails indirect costs in the bus/corridor equilibrium but not in the Solow equilibriumm whether the shadow rent on land in road use is higher or lower in the the Solow equilibrium than in the bus/corridor equilibrium is *a priori* ambiguous, and in the example the two shadow rent functions "intersect" twice. The sequal to this paper will compare the two equilibria with residential land use endogenous, and with and without the allocation of land between roads and housing being optimized.

4.3 Social Optimum

The bus discrete approximation of the corridor problem worked well in the no-toll equilibrium problem in the sense that it generated qualitatively the same solution as the pure corridor problem. It is not without its problems, however. The equation of continuity links congestion at a location to congestion both upstream and downstream. The bus approximation severs this link, with the speed of a bus unrelated to the situation of buses either upstream or downstream. This creates a difficulty in the bus approximation of the social optimum corridor problem.

To illustrate, write down the social optimum problem, on the assumption that in the social optimum all buses except the last are strictly early, which would be true if the qualitative departure pattern were a horn pattern, as in the no-toll equilibrium:

$$\begin{aligned} \begin{Bmatrix} n_j^i \end{Bmatrix} \\ \begin{Bmatrix} P_j^i \end{Bmatrix} \end{aligned} \min \alpha \sum_{j=1}^{J} \sum_{i=1}^{I} \quad P_j^i \tau_j^i + \beta \sum_{j=1}^{J} \left(t^* - \left(t_1 + (j-1)\, s + \sum_{i=1}^{I} \tau_j^i \right) \right) P_j^I$$

$$\begin{aligned} s.t. \quad i) \qquad & \sum_{j=1}^{J} n_j^i = N^i, \quad i = 1, ..., I \\ ii) \qquad & n_j^i \geq 0, \quad i = 1, ..., I, \quad j - 1, ..., J18 \\ iii) \qquad & P_j^i = \sum_{i=1}^{i} n_j^i, \quad i = 1, ..., I, \quad j = 1, ..., J \end{aligned} \tag{22}$$

$$iv) \qquad t^* \geq t_1 + (J-1)\,s + \sum_{i=1}^{I} \tau_J^i,$$

where t_1 is the time of the first bus to depart and $\tau_j^i = \tau_j^i\left(P_j^i\right)$. The solution minimizes the sum of the total cost of travel time plus the total cost of time early, subject to the constraints that: *i*) everyone commute; *ii*) $\left\{n_j^i\right\}$ be non-negative; *iii*) the number of passengers on the bus between stops i and $i+1$ equal the number of passengers who have boarded the bus at stops 1 through i; and *iv*) the J^{th} bus cannot arrive late. Since early buses may be empty, the choice of J is arbitrary as long as it does not constrain the solution, which is assumed.

Let Ω^i denote the shadow price on constraint *i*) for bus stop i , which equals the marginal social cost of an extra passenger at bus stop i. Solving the social optimum problem and substituting out the Lagrange multipliers yields

$$\Omega^i = \begin{cases} \sum_{i'=i}^{I} \left\{\alpha\left(\tau_j^{i'} + P_j^{i'}\frac{\partial \tau_j^i}{\partial P_j^{i'}}\right) - \beta\frac{\partial \tau_j^i}{\partial P_j^{i'}}P_j^I\right\} + \beta\left(t^* - (t_1 + (j-1)\,s) - \sum_{i'=i}^{I}\tau_j^{i'}\right) \\ for j = 1, ..., J-1 \\ \\ \sum_{i'=i}^{I} \left\{\alpha\left(\tau_J^{i'} + P_J^{i'}\frac{\partial \tau_J^i}{\partial P_J^{i'}}\right) - \beta\frac{\partial \tau_J^i}{\partial P_J^{i'}}P_J^I\right\} + \beta\left(\sum_{i'=i}^{I} N^{i'}\right)\left(\sum_{i'=i}^{I}\frac{\partial \tau_J^{i'}}{\partial P_J^{i'}}\right) \\ for j = J \end{cases} \quad 19 \quad (23)$$

Since, by the assumption, buses $j = 1, ...J-1$ are strictly early, the marginal cost of carrying the extra passenger at stop i on any bus except the last is the user cost $\left(\sum_{i'=i}^{I}\alpha\tau_j^{i'} + \beta\left(t^* - (t_1 + (j-1)\,s) - \sum_{i'=i}^{I}\tau_j^{i'}\right)\right)$ plus the congestion externality cost $\left(\sum_{i'=i}^{I}\left\{\alpha P_j^{i'}\frac{\partial \tau_j^{i'}}{\partial P_j^{i'}} - \beta\frac{\partial \tau_j^{i'}}{\partial P_j^{i'}}P_j^I\right\}\right)$. The congestion externality cost equals the increase in total travel time costs the extra passenger imposes on others minus the decrease in total time early costs due to later arrival of the bus deriving from the increase in travel time caused by the extra passenger. The expression for the user cost of bus J is the same as that for the previous buses, as is the travel time component of the congestion externality. However, there is an additional cost. To satisfy the condition that the last bus arrive no later than t^*, the starting times of *all* buses have to brought forward by $\sum_{i'=i}^{I}\frac{\partial \tau_J^{i'}}{\partial P_J^{i'}}$. The result — that an additional passenger causes additional total time early costs only when put on the last bus — is entirely an artifact of the assumed bus technology.

This problem is dealt with in *ad hoc* way. A pseudo social optimum is solved for in which the toll is set equal to the travel time externality cost. With the Vickrey congestion cost function, the toll would then be $\pi_j^i = \gamma^\alpha \left(T_j^i - F^i\right)$ when T_j^i is the time to travel from bus stop i to the depot on bus j, in excess of the free flow travel cost, F^i. Accordingly,

$$\begin{aligned} p^i &= \alpha T_j^i + \beta\left(t^* - \left(t_j^i + T_j^i\right)\right) + \pi_j^i \\ &= \alpha T_j^i + \beta\left(t^* - \left(t_j^i + T_j^i\right)\right) + \gamma\alpha\left(T_j^i - F^i\right) \\ &= \alpha F^i + \beta\left(t^* - \left(t_j^i + F^i\right)\right) + \left(\left(1+\gamma\right)\alpha - \beta\right)\left(T_j^i - F^i\right) 20 \end{aligned} \tag{24}$$

if $n_j^i > 0$. Comparing ρ^i on buses j and $j+1$ on the assumption that both $n_j^i > 0$ and $n_{j+1}^i > 0$ yields

$$\begin{aligned} & \beta\left(t^* - \left(t_j^i + F^i\right)\right) + \left(\left(1+\gamma\right)\alpha + \beta\right)\left(T_j^i - F^i\right) \\ = \; & \beta\left(t^* - \left(t_{j+1}^i + F^i\right)\right) + \left(\left(1+\gamma\right)\alpha + \beta\right)\left(T_{j+1}^i - F^i\right) \\ & or \\ & \beta\left(t_{j+1}^i - t_j^i\right) = \left(\left(1-\gamma\right)\alpha + \beta\right)\left(T_{j+1}^i - T_j^i\right) 21 \end{aligned} \tag{25}$$

Assume that the departure pattern is horn-shaped. Then the travel time from the boundary of the city to bus stop i is the same for both buses, so that $t_{j+1}^i - t_j^i = s$. and $T_{j+1}^i - T_j^i = \frac{\beta s}{(1-\gamma)\alpha+\beta}$. Let $j(i)$ denote the first bus which picks up passengers at stop i. Then for $j > j(1)$

$$T_{j+1}^1 - T_j^1 = \frac{\beta s}{\left(1-\gamma\right)\alpha + \beta} 22 \tag{26}$$

The departure pattern is that which would occur in the no-toll equilibrium if α were replaced by $(1-\gamma)\alpha$.The corresponding equilibrium is termed the *pseudo social optimum.* [8].

Table 2 displays the base case pseudo social optimum

[8]Perhaps the pseudo social optimum can be derived by imposing the additional constraint on (18) that the headway between buses can be no smaller than s : $\sum_{i'=1}^{i} \tau_{j+1}^{i'} \geq \sum_{i'=1}^{i} \tau_j^{i'}$ for all $j \leq J-1$.

	1	2	3	4	5	6	7	8	
N^i	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	
w^i	200.0	200.0	200.0	200.0	200.0	200.0	200.0	200.0	
n_1^i	$16.\dot{6}$	$16.\dot{6}$	1.91	0.0	0.0	0.0	0.0	0.0	35.2
n_2^i	$16.\dot{6}$	$16.\dot{6}$	19.62	20.0	1.52	0.0	0.0	0.0	74.5
n_3^i	$16.\dot{6}$	$16.\dot{6}$	19.62	20.0	24.62	11.25	0.0	0.0	108.82
n_4^i	$16.\dot{6}$	$16.\dot{6}$	19.62	20.0	24.62	29.58	18.13	0.0	145.25
n_5^i	$16.\dot{6}$	$16.\dot{6}$	19.62	20.0	24.62	29.58	40.94	23.79	191.88
n_6^i	$16.\dot{6}$	$16.\dot{6}$	19.62	20.0	24.62	29.58	40.94	76.21	244.30
p^i	9.971	9.064	8.139	7.176	6.157	5.042	3.779	2.243	
ρ^i	0.000	0.023	0.115	0.713	0.893	0.996	1.752	3.301	
$\widehat{p}^i$	7.500	7.176	6.774	6.249	5.550	4.626	3.408	1.900	
$\widehat{\rho}^i$	0.008	0.066	0.223	0.529	1.034	1.787	2.837	4.235	

$T_6^1 = 0.554 \quad t^* - t_1^1 = 1.054 \quad TCTC = 1162$

$\theta = 0.387 \quad TC = 2830 \quad TEC = 660$

$R = 2324 \quad TFC = 1008$

Notes: See Table 1 for parameter values

4.4 Comparisons

This subsection will make a variety of comparisons. The first is to compare the non-toll equilibrium and the pseudo social optimum for the bus/corridor problem.

> Departure pattern. Total trip cost. Proportion of total trip costs in excess of free flow travel time costs saved by imposition of toll

In the Solow formulation, in contrast, with fixed land use there are no efficiency gains to be achieved from congestion tolling since in that formulation land use is the only possible margin of adjustment to tolling. The next section will treat variable land. Then it will be of interest to investigate, for the bus/corridor by what proposition the efficiency gains from congestion tolling are increased when land use is variable.

The next comparison to be made is between the pattern of the shadow rent on the land in road use for the bus/corridor problem with and without the optimal toll, and for the Solow formulation with $\{N^i\}$ freed. The shadow rent in land use is computed as the reduction in total trip costs from increasing w^i by one unit, which we interpret as adding a unit area of land to the road, at the various bus stops. The results are reached in Table 3.

INSERT TABLE 3 HERE

The shadow rent may be separated into two components, the direct benefit which derives from adjustment in the departure pattern induced by the increase in road width. For the Solow formulation, there is no adjustment in departure pattern, and consequently indirect cost is zero. For the bus/tolling case, since the departure pattern is optimized, by the Envelope Theorem the ratio of indirect cost to direct benefit is zero for a infinitesimal expansion of road width, and "small" for a small proportional change in road width. For the bus/no-tolling case, the indirect costs are of the same order of magnitude as the direct benefits and could exceed these. In the above example, . . .TO BE CONTINUED

5 The Corridor Problem with Variable Land Use: Buses

5.1 Introduction

There are two interesting ways in which land use may be varied. The first is to endogenize land use, the second is to optimize the allocation of land between residential and roads. Kanemoto (____) and Arnott (____) provided through analyses of the corridor problem with variable land use in a monocentric city with the Solow formulation of congestion. They focused on two issues, first on the relationship between the market rent and shadow rent on residential land at different locations, when congestion is unpriced, and second on the characteristics of the misallocation of land between