This paper shows that the inefficiency of fiscal decentralization in the presence of spillovers, a main tenet of the decentralization literature, is overturned in a particular transportation context. In a monocentric city where road (bridge) capacity is financed by budget-balancing user fees, decentralized capacity choices (made by individual zones within the city) generate the social optimum despite the presence of spillovers. Optimality also requires the correct population distribution across the city’s zones, conditional on bridge capacities. This outcome is achieved because the user fees function as optimal congestion tolls, a result that follows from the famous self-financing theorem of transportation economics.

Keywords: congestion, spillovers, decentralization

JEL Codes: H71, R41

I. INTRODUCTION

Whether fiscal decentralization is optimal depends on the presence of interjurisdictional spillovers in policy decisions. As argued by Oates (1972), local policy choices, which beneficially allow fulfillment of local preferences, are nevertheless inefficient when significant public-sector spillovers exist. The reason is that the external benefits (or costs) of a policy will be ignored by local decision-makers. Under these circumstances, centralization of public-sector decisions may be desirable. This argument, which was recently reformulated in a precise fashion by Besley and Coate (2003), helped to spawn a vast literature on the advantages and disadvantages of decentralization.

Subsidization of public goods by the central government is one remedy for the inefficiency generated by spillovers. With subsidies properly set, the provision of public

1 Their analysis was offered as a prelude to development of a different model with features at variance with the Oates tradition.
goods expands to a level justified by the existing spillovers. While this remedy was discussed by Oates (1972) and many subsequent authors (e.g., Conley and Dix, 1999), Wellisch (1993) offered a different approach, showing that a particular institutional structure eliminates the inefficiency problem. Building on work by Myers (1990), Wellisch demonstrated that, if localities are able to make voluntary horizontal transfers to other jurisdictions, and if they recognize how their actions affect the distribution of the mobile consumer population, then an equilibrium with spillovers is efficient. Efficiency also arises in the model of Ogawa and Wildasin (2009), where the spillover is due to the pollution generated by the mobile capital used in a jurisdiction, which is subject to a tax that finances a public good. While the capital flight caused by a high local tax confers a positive tax-base externality on other jurisdictions, it also generates a loss by raising their pollution levels (they trade a pollution spillover from the original jurisdiction for more intense locally-generated pollution). However, these distortions affecting the capital-tax choice exactly cancel, making the equilibrium tax rates efficient.

Another route to attainment of efficiency with spillovers comes from recognizing that some public goods are generated by public facilities for which user fees can be levied. Since these fees are paid both by local residents and by outsiders using the facility, they lead to cost sharing that in turn prompts a beneficial expansion of the facility’s size relative to the case where costs cannot be recouped from outsiders, which tends to eliminate the inefficiency. Despite this simple insight, the literature on fiscal decentralization has left the user-fee approach to handling spillovers mostly unexplored. As a result, the conditions under which the approach might be fully successful are not generally known.

Transportation infrastructure represents a public good where spillovers exist, with consumers from one locality using the infrastructure of another, and where user fees, possibly in the form of road tolls, are feasible. Recently, several papers have explored the efficiency of decentralized infrastructure investment and pricing in the transportation context, offering greater understanding of the role of user fees as a possible remedy for the spillover problem. Following the spirit of earlier work by De Borger, Proost, and Van Dender (2005) and De Borger, Dunkerley, and Proost (2007), De Borger and Proost (2013) develop a transportation model with spillovers that fits neatly within the Oates tradition. Moreover, given the road-transportation context, where consumers in one locality drive on both their own congested roads and those of a neighboring jurisdiction, user fees are a natural financing method. The paper analyzes decentralized outcomes under a variety of different assumptions, focusing mostly on the case where user fees can generate a profit over and above the infrastructure cost (which is then redistributed to local residents). However, in the case where balanced-budget user fees must be levied, the paper shows that the decentralized equilibrium is efficient under a common set of assumptions on the transportation technology.2

Kidokoro (2014) analyzes a model where user fees can be differentiated between local and outside users, allowing the locality to inefficiently exploit market power over the latter group. Efficiency is obtained, however, if the outside user fee is combined with a lump-sum charge designed to extract all consumer surplus. Mun and Nakagawa (2010) analyze a transportation-spillover model with a different structure, considering a number of different financing schemes, none of which turn out to be efficient.

---

2 Kidokoro (2014) analyzes a model where user fees can be differentiated between local and outside users, allowing the locality to inefficiently exploit market power over the latter group. Efficiency is obtained, however, if the outside user fee is combined with a lump-sum charge designed to extract all consumer surplus. Mun and Nakagawa (2010) analyze a transportation-spillover model with a different structure, considering a number of different financing schemes, none of which turn out to be efficient.
The purpose of the present paper is to derive a parallel efficiency result in a somewhat different transportation context. The analysis focuses on a congested, monocentric city, building on the framework of Brueckner and Helsley (2011). Analysis of congested cities has a long tradition in urban economics, and a major lesson of this literature is that, in the absence of congestion pricing, cities overexpand, creating inefficient urban sprawl. While optimal congestion tolls, imposed by a social planner, can eliminate this tendency, previous analyses have not investigated whether such tolls would be imposed under decentralized decision-making.

The paper addresses this question in a model where spatial spillovers generate another potential distortion that threatens efficiency. In the model, the city is composed of separate jurisdictions, called “zones,” which are connected by a sequence of locally controlled, congestible bridges that lead to the central business district (CBD). A spillover exists in bridge-capacity choices since bridges are used by residents of other zones, and the congestion externality also distorts individual decisions. The conclusion of the analysis is that, despite these potential distortions, decentralized investment and pricing is efficient, leading to an urban equilibrium that coincides with the social optimum. In other words, bridge capacities are chosen efficiently, and optimal congestion tolls are charged, ensuring an efficient distribution of the population conditional on bridge capacities.

To understand the details underlying these conclusions, consider Figure 1, which shows the configuration of the city, extending the setup used by Brueckner and Helsley (2011). The central zone at the left contains the CBD along with residential land, while the “midcity” and suburban zones lie to the right. The midcity bridge, which is controlled by the midcity zone, connects that zone to the center, while the suburban bridge (again controlled locally) connects the suburbs to the midcity zone. Suburban residents must cross both bridges to reach the CBD, while midcity residents need only cross their own bridge (central residents are not bridge users). Intrazone travel costs are zero, so that the only commuting costs are incurred on the congested bridges. Travel costs are high on these bridges, generating congestion.

![Figure 1](image-url)

City Configuration

---

demand is completely inelastic, so that bridge traffic levels are influenced only by the endogenous distribution of the population across the city’s zones.

Bridge capacity is financed by user fees. In addition, capacity is produced with constant returns to scale, and bridge congestion is homogenous of degree zero in traffic volume and capacity, conditions (also invoked by De Borger and Proost, 2013) that yield the famous self-financing theorem of transportation economics. This theorem, due to Mohring and Harwitz (1962), states that congestion tolls exactly cover the cost of an optimal-size road or bridge (Small and Verhoef, 2007). The governments of the midcity and suburban zones choose their individual bridge capacities to maximize resident utilities, and the result is an efficient equilibrium, which yields optimal bridge capacities and an optimal distribution of population across the city’s zones.

This outcome is surprising at first because spillovers, which usually lead to inefficiency, are present, as are congestion externalities. In particular, the midcity’s capacity choice affects congestion on its bridge and thus the travel costs of suburban residents; midcity and suburban commuters impose congestion costs on one another on the midcity bridge, and suburban residents impose congestion costs on one another on the suburban bridge. While congestion should be addressed by tolls, the capacity spillover (which tends to make the chosen midcity bridge too small) would appear to require a different remedy.

The use of budget-balancing user fees simultaneously addresses both problems. Because the midcity fees are paid by both midcity and suburban commuters, capacity costs for the midcity bridge are shared with outsiders, encouraging the zone’s residents to expand its size, perfectly counteracting the omission of suburban benefits in their capacity decision. While this remedial effect was noted above, a further crucial fact is that, under the conditions of the self-financing theorem, the user fees charged on the two bridges end up functioning as optimal congestion tolls, correcting the congestion externalities. In other words, since user-fee revenue matches capacity cost, and since this cost would also be exactly covered by revenue from the optimal toll, it follows that the user fees on the two bridges coincide with optimal tolls.

Thus, as a result of the confluence of these disparate factors, decentralized decisions made by individual local governments are efficient in a transportation context despite the presence of spillovers and congestion externalities, showing circumstances under which user fees can lead to a desirable decentralization outcome. This finding is new in a monocentric-city setting, although it matches De Borger and Proost’s (2013) result from a model without an explicit spatial structure and land market, where equilibria are determined by majority voting in the presence of multiple interest groups. But given the importance of the efficiency finding, its demonstration in a variety of contexts is worthwhile.

The plan of the paper is as follows. Section II develops the model and derives the main efficiency result. Section III shows how relaxing the key assumptions of the model overturns efficiency. Section IV considers centralized decisions, and Section V offers conclusions.

---

4 The present work was inspired by De Borger’s brief presentation of the 2013 paper in a conference setting, which made clear the usefulness of analyzing decentralization in a transportation context. My later, fuller reading of their paper showed that the authors had already derived the current efficiency result as part of a much broader analysis.

5 For analysis of decentralized transportation investment in a monocentric city that relies on a much different framework, see Ferguson (2013).
II. CONSTRUCTION AND ANALYSIS OF THE MODEL

The variables in the model are as follows. Individual land consumption is denoted \( q_i \), with consumption in the suburban, midcity, and central zones denoted by \( q_s \), \( q_m \), and \( q_c \). The opportunity cost of land (the agricultural rent) equals \( r \). The non-land composite good, which is the numeraire, is denoted \( e \), and consumption levels in the three zones are given by \( e_i, i = s, m, c \). To simplify the exposition, preferences are assumed to take the quasi-linear form \( e + V(q) \), where \( V' > 0, V'' < 0 \). The paper’s main results continue to hold with a more general specification of preferences. The populations of the three zones are \( n_i, i = s, m, c \), and they must sum to the city’s fixed overall population \( N \), so that

\[
(1) \quad n_s + n_m + n_c = N.
\]

The capacities of the suburban and midcity bridges are denoted by \( k_s \) and \( k_m \). The costs of providing these capacities are given by \( B(k_i) = 0, i = s, m \), where \( B' > 0 \). With constant returns to scale, \( B'' = 0 \), while \( B'' > (<)0 \) holds under decreasing (increasing) returns. The time cost of crossing a bridge depends on the level of congestion, which is determined by the traffic volume and the bridge’s capacity.\(^6\) With the traffic volume on the suburban bridge equal to \( n_s \), the time cost of crossing this bridge is \( T(n_s, k_s) \), where the partial derivatives of \( T \) satisfy \( T_1 > 0, T_2 < 0 \). With traffic on the midcity bridge equal to \( n_s + n_m \), the time cost of crossing it equals \( T(n_s + n_m, k_m) \). Under the conditions of the self-financing theorem, \( T \) is homogeneous of degree zero, being a function only of the volume/capacity ratio. Finally, travel costs within each zone are zero.

A. The Social Optimum

The planner’s goal is to minimize the city’s resource consumption while generating a fixed common utility level \( u \) for the residents in each of the zones. Since it is inefficient to leave vacant land inside the city, the central and midcity zones will be fully occupied, with vacant land found only in the suburban zone. Normalizing the individual land areas of the two inner zones to 1, the constraints

\[
(2) \quad n_s q_s = 1, n_m q_m = 1
\]

must be satisfied. Imposing these two constraints, the utility constraints, and the overall population constraint, the Lagrangian expression for the planner’s problem is

\[
(3) \quad L = n_s e_s + n_m e_m + n c e_c + r (n_s q_s + n_m q_m + n_c q_c) \\
+ n T(n_s, k_s) + (n_s + n_m) T(n_s + n_m, k_m) + B(k_s) + B(k_m) \\
+ \sum_{i=s,m,c} \lambda_i (e_i + V(q_i) - u) + \gamma (n_s + n_m + n_c - N) + \sum_{i=s,m,c} \theta (n_i q_i - 1).
\]

\(^6\) The time used in commuting is assumed to reduce work hours and output, thus yielding a pecuniary cost.
The terms up to the first summation sign in (3) capture total resource consumption, which equals total consumption of $e$, plus the opportunity cost of the urbanized land, plus bridge crossing costs, plus the cost of bridge capacities. The various constraints, with Lagrange multipliers appended, appear in the last line of (3).

Differentiating (3) with respect to $e_i$ yields $n_i = \lambda_i$, $i = s, m, c$. Using these equalities and differentiating (1) with respect to $q_i$, $i = s, m, c$, yields the first-order conditions

(4) \[ V'(q_i) = \bar{r} \]

(5) \[ V'(q_i) = \bar{r} + \theta_i, \quad i = m, c. \]

These conditions equate the marginal utility of land consumption to the shadow price of land in a zone, with the suburban shadow price equal to the opportunity cost of land $\bar{r}$ and the $\theta$’s equal to the price premia in the inner zones.

The first-order conditions for bridge capacities, $k_i$, $i = s, m$ are

(6) \[ B'(k_s) + n_s T_z(n_s, k_s) = 0 \]

(7) \[ B'(k_m) + (n_s + n_m) T_z(n_s + n_m, k_m) = 0, \]

which indicate that the total time-cost savings from an increase in capacity (the negative of the second term) equals marginal capacity cost.

The first-order conditions for $n_c$ and $n_m$ are

(8) \[ e_c + (\bar{r} + \theta_c) q_c = -\gamma \]

(9) \[ e_m + (\bar{r} + \theta_m) q_m + T(n_s + n_m, k_m) + (n_s + n_m) T_1(n_s + n_m, k_m) = -\gamma. \]

These conditions imply that the resource costs of adding an extra person to the central and midcity zones should be the same and equal to $-\gamma$. These costs include the individual’s $e$ consumption, the cost of her land (evaluated at the shadow price), and the extra commuting cost generated by her presence. While this cost is zero for someone added to the center, a person added to the midcity zone incurs her own cost $T(•)$ (in (9)) while raising time costs by $T_1$ for each of the other $n_s + n_m$ commuters on the midcity bridge (aggregating yields the last term in (9)).

Equations (8) and (9) can be combined by substituting $e_i = u - V(q_i)$, $i = s, m$, into the LHS expressions and then equating these expressions to eliminate $\gamma$. Multiplying through by $-1$, the result is

(10) \[ V(q_c) - (\bar{r} + \theta_c) q_c = V(q_m) - (\bar{r} + \theta_m) q_m - (n_s + n_m) T_1(n_s + n_m, k_m) - T(n_s + n_m, k_m). \]

The first-order condition for $n_s$ is $e_s + \bar{r} q_s + T^s + n_s T^s_1 + T^m + (n_s + n_m) T^m_1 = -\gamma$, where the $T$ superscripts indicate that the functions are evaluated in zones $m$ and $s$. The LHS
expression gives the cost of adding a resident to the suburban zone, which generates
time and congestion costs on both bridges. This condition is transformed by eliminating
$e_s$, multiplying through by $-1$, and setting the result equal to the RHS of (10) (which
equals $\gamma - u$). The terms $-T_m - (n_s + n_m)T_1^m$ and $u$ are common to both expressions and
thus cancel, yielding the condition

$$(11) \ V(q_m) - (\bar{r} + \theta_m)q_m = V(q_s) - \bar{r}q_s - n_sT_1(n_s, k) - T(n_s, k).$$

Satisfaction of conditions (10) and (11) implies that the resource costs of adding a
person to the central, midcity or suburban zones are all equal.

The 10 conditions consisting of (1), (2), (4–7), (10), and (11) determine the socially
optimal values of the 10 variables $q_c, q_m, q_s, n_s, n_m, n_c, k_s, k_m, \theta_m$, and $\theta_c$. The $e$
consumption levels in the three zones can be recovered from the utility constraints. Note that the only
effect of a change in the parametric utility level $u$ is to change the $e$ values. Thus, unlike
in the case where income effects are present, the values of the 10 variables above are the
same in all social optima (that is, regardless of the magnitude of the fixed utility level $u$).

**B. The Decentralized Equilibrium**

Turning to a characterization of the decentralized equilibrium, the urban residents
earn a maximal income of $y$ from employment at the CBD, which is reduced by the loss
of work time from congested bridge crossings. As explained in the introduction, each
zone is a separate jurisdiction, with the governments of the midcity and suburban zones
having the power to determine the capacities of their respective bridges. In addition, the
two zones are able to levy user fees on commuters crossing their bridges. Users of the
suburban bridge consist only of the zone’s residents, so that suburban user-fee revenue
is internal to that zone. But users of the midcity bridge consist of both suburban and
midcity residents, so that some of the midcity’s user-fee revenue comes from suburban
residents. Finally, each jurisdiction must satisfy a balanced-budget requirement, with
user-fee revenue exactly covering the capacity cost of its bridge.

Given these assumptions and letting $r_m$ denote midcity land rent, the budget constraint
of a midcity resident is

$$(12) \ e_m = y - r_m q_m - \frac{B(k_m)}{n_s + n_m} - T(n_s + n_m, k_m),$$

where the third term on the RHS is the budget-balancing user fee (capacity cost divided
by the number of users of the midcity bridge). The suburban resident’s budget con
straint is

$$(13) \ e_s = y - \bar{r}q_s - \frac{B(k_s)}{n_s} - \frac{B(k_m)}{n_s + n_m} - T(n_s, k_s) - T(n_s + n_m, k_m).$$

Note that a suburban resident incurs time costs on two bridges and pays user fees on
each one (the third term on the RHS of (13) is the suburban fee). Note also that the
land rent paid by this resident equals $\bar{r}$, the value of agricultural land. Finally, since
residents of the central zone incur no commuting cost, their budget constraint is simply
\[ e_c = y - r_c q_c, \]
where \( r_c \) denotes central land rent.

Zone residents choose land consumption levels to maximize their utilities, \( e_i + V(q_i), \)
\( i = s, m, c. \) Using the budget constraints, the first-order conditions are

\[ (14) \quad V'(q_s) = \bar{r} \]
\[ (15) \quad V'(q_i) = r_i, \quad i = m, c. \]

In addition, zone governments, acting on behalf of the suburban and midcity residents,
choose the capacities of their bridges. In doing so, each government views the other
zone’s capacity choice and the zone populations as parametric, even though these popu-
lations are ultimately determined by the capacity choices. Using the budget constraints
(12) and (13), the resulting first-order conditions for \( k_s \) and \( k_m \) are

\[ (16) \quad \frac{B'(k_s)}{n_s} + T_z(n_s, k_s) = 0, \]
\[ (17) \quad \frac{B'(k_m)}{n_m + n_s} + T_z(n_s + n_m, k_m) = 0. \]

These conditions state that the marginal increase in the user fee due to a capacity
expansion equals the resulting saving in individual time cost. Conditions (16) and (17)
are equivalent to the planning conditions (6) and (7), resulting from division of each
planning condition by bridge traffic \( n_i \) or \( n_i + n_m \).

As usual in urban models, additional equilibrium conditions require equalization of
utilities across zones. In other words, \( e_i + V(q_i) = e_m + V(q_m) = e_s + V(q_s) \) must hold.

Using the budget constraints to eliminate the \( e \)'s in the first equality and cancelling the
\( y \)'s that appear on both sides, the equality requires

\[ (18) \quad V(q_s) - r_s q_s = V(q_m) - r_m q_m - \frac{B(k_m)}{n_s + n_m} - T(n_s + n_m, k_m). \]

Turning to the second of the previous equalities, the budget constraints are again used
to eliminate the \( e \)'s, and both the common \( y \)'s and the common terms \( B(k_m)/(n_s + n_m) + T^m \)
are cancelled (see (12) and (13)), yielding

\[ (19) \quad V(q_m) - r_m q_m = V(q_s) - \bar{r} q_s - \frac{B(k_s)}{n_s} - T(n_s, k_s). \]

The equilibrium values of \( q_c, q_m, q_s, n_s, n_m, e_c, k_c, k_m, r_m, r_c \) are determined by 10 equi-
librium conditions: the population and land area constraints in (1) and (2), the first-order
conditions (14–17), and the utility-equalization conditions (18) and (19). Recognizing
that the land shadow prices $\bar{r} + \theta_c$ and $\bar{r} + \theta_m$ and the rents $r_c$ and $r_m$ are equivalent, the equilibrium and optimality conditions can be compared.

C. Comparing the Equilibrium and the Optimum

The constraints (1) and (2) are common to the equilibrium and optimality conditions, and the first-order conditions for the $q$’s are also the same in the two cases. As noted above, the capacity-choice conditions (16) and (17) are also the same as the corresponding planning conditions, and this equivalence indicates that the potential distortion from the capacity spillover associated with the midcity bridge is corrected in the equilibrium. To see how, recall that the distortion arises because midcity residents have no incentive to consider the benefits to suburban residents in choosing their bridge capacity. However, since user-fee financing means that the midcity bridge is partly paid for by outsiders (suburban residents), midcity decision-makers are encouraged to expand its capacity. This incentive exactly cancels the tendency to underprovide capacity due to the spillover, yielding an optimal outcome. Thus, reliance on user fees is crucial in generating an optimal bridge capacity.

In comparing the planning and equilibrium conditions, the remaining comparison is between the population-allocation conditions (10–11) and the equal-utility conditions (18–19). The conditions are different, with the term $(n_s + n_m)T_1^m$ in (10) replaced by $B(k_m)/(n_s + n_m)$ in (18) and an analogous difference seen in (11) and (19). However, under the assumptions of the self-financing theorem, the conditions are the same. These assumptions are constant returns in provision of bridge capacity, implying $B(k_m) = \beta k_m$, and zero-degree homogeneity of $T$, which implies $(n_s + n_m)T_1^m + k_m T_2^m = 0$ or $T_2^m = -[(n_s + n_m)/k_m]T_1^m$. Using this latter relationship to replace $T_2^m$ in (7), the condition becomes

\[(20)\quad B'(k_m) - \frac{(n_s + n_m)^2}{k_m}T_1(n_s + n_m, k_m) = 0.\]

Noting $B' = \beta$ and rearranging, (20) reduces to

\[(21)\quad \frac{B(k_m)}{n_s + n_m} \equiv \frac{\beta k_m}{n_s + n_m} = (n_s + n_m)T_1(n_s + n_m, k_m).\]

Therefore, $(n_s + n_m)T_1^m$ in the planning condition (10) can be replaced by $B(k_m)/(n_s + n_m)$, making that condition the same as the equilibrium condition (18). Thus, when capacity is chosen optimally, the midcity user fee in (21) equals the optimal congestion toll, given by the expression on the RHS that captures the congestion damage from an extra midcity-bridge commuter. For the city’s population to be optimally distributed, these commuters must face such a toll. The same argument shows that the suburban bridge’s user fee also equals the optimal toll, establishing that (11) and (19) are the same and thus that the equilibrium and social optimum coincide. Note that, after multiplying through by $n_s + n_m$, (21) implies that capacity cost equals congestion toll-revenue (as specified in the self-financing theorem).
Thus, by relying on budget-balancing user fees, the midcity residents choose their bridge capacity in a socially optimal fashion despite the presence of a spillover. Given the self-financing theorem, the user fee associated with this optimal capacity then coincides with the optimal congestion toll, ensuring a proper distribution of the population across zones. Note that with a capacity spillover absent, the only distortion affecting suburban choices is congestion, which is handled correctly by the optimal toll.

Summarizing yields

**Proposition 1.** Suppose that the conditions of the self-financing theorem (constant returns in provision of capacity and zero-degree homogeneity of congestion costs) are satisfied. Then, provided that financing relies on budget-balancing user fees, decentralized choice of road (bridge) capacities in a multijurisdictional city with spillovers leads to an efficient equilibrium.

## III. RELAXING THE KEY ASSUMPTIONS

### A. Relaxing the Technological and User-Cost Assumptions

Efficiency disappears under alternate assumptions. Consider first a different bridge financing arrangement, under which zone governments rely on local tax revenue to pay for bridge capacity, which is then used by outsiders free of charge. In this case, the user fee $B(k_m)/(n_s + n_m)$ is replaced by a midcity head tax of $B(k_m)/n_m$ in (12) and by zero in (13), indicating free use of the midcity bridge by suburban residents. The first-order condition for decentralized choice of $k_m$ becomes

$$\frac{B'(k_m)}{n_m} + T_z(n_s + n_m, k_m) = 0,$$

which differs from the optimality condition (7). The absence of cost sharing, which reduces the factor dividing $B'(k_m)$ and thus raises the per capita cost of bridge expansion, inefficiently reduces the incentive of midcity residents to do so.

While this inefficiency means that the equilibrium leads to higher resource consumption than the social optimum (see (3)), the difference in the values of the model’s variables between the new equilibrium and the previous efficient one is ambiguous in general. But analysis that is available on request shows that if $\eta^m$ is close to zero, indicating high price sensitivity of the demand for land, then the new financing scheme reduces $k_m$ relative to its value in the previous efficient equilibrium, an outcome that conforms to intuition. In addition, population shifts from the midcity to the central zone as some commuters seek to avoid the smaller midcity bridge, with $n_c$ rising and $n_m$ falling relative to the efficient equilibrium ($q_c$ and $q_m$ move in directions opposite to the zone populations). However, the change in $n_s$, and thus in the city’s overall spatial size, remains ambiguous.7

Suppose instead that user-fee financing is retained but that bridge capacity is produced with non-constant returns. Then, toll revenue no longer equals the cost of optimal

---

7 Note that $q_s$ (which is tied to $\tilde{r}$) and $n_s/k_s$ remain constant, being unaffected by the financing used for the midcity bridge.
capacity, and budget-balancing user fees will not coincide with optimal congestion tolls, yielding an inefficient equilibrium. To see this conclusion, suppose that capacity costs are given by $\beta k^\alpha$, with $\alpha \neq 1$. Then $B'$ in the capacity condition (7) is replaced by $\alpha \beta k_m^\alpha$. With $(n_s + n_m)T_2$ again equal to $-[(n_s + n_m)^2/k_m]T_1$, rearrangement of (7) yields

$$
\alpha \beta k_m^\alpha \frac{n_s}{n_s + n_m} = \frac{\alpha B(k_m)}{n_s + n_m}.
$$

Substitution in (10) then yields a condition that is no longer the same as the equilibrium condition (18), a consequence of the new $\alpha$ factor. This conclusion also applies to (11) and (19). The equilibrium is thus inefficient, leading to higher than optimal resource consumption, but the directions in which the individual variables diverge from the optimal values are mostly ambiguous, although a few comparisons can be derived for special cases. Relaxing the assumption of zero-degree homogeneity of the $T$ function can similarly be shown to eliminate efficiency.8

B. Introducing Income Heterogeneity

While relaxing the explicit assumptions of Proposition 1 overturns the efficiency result, the proposition also depends on a further implicit assumption, whose relaxation has the same effect. In particular, while the analysis has assumed that consumers are homogeneous, introducing income heterogeneity can overturn Proposition 1. In particular, suppose the city has two income groups, with $y_H$ and $y_L < y_H$ denoting the incomes of the high- and low-income groups, respectively. With quasi-linear preferences, this income difference by itself has no effect on the analysis, but income differences will translate into time-cost differences and thus into group-specific $T$ functions, with $T^H(n, k) > T^L(n, k)$ holding for common traffic and capacity levels $n$ and $k$ (reflecting higher time costs for the $H$ group).

Although the locational pattern of high- and low-income groups in a monocentric city is well known to be ambiguous, suppose that the $H$ group lives farther from the CBD than the $L$ group. Moreover, suppose for simplicity that, in equilibrium, the $H$ group exactly fits in the suburban zone and the $L$ group exactly fits in the two inner zones. In this case, the condition for optimal capacity of the midcity bridge (previously (7)) becomes

$$
B'(k_m) + n_s T^H_2(n_s + n_m, k_m) + n_m T^L_2(n_s + n_m, k_m) = 0.
$$

In addition, $T_2$ in (16) acquires an $H$ superscript, and the $T$ and $T_1$ terms in (10) are replaced by separate terms involving the $T^H$ and $T^L$ functions.

To see the inefficiency of the equilibrium, it is sufficient to focus on the capacity choice for the midcity bridge. In the equilibrium condition (17), the derivative $T_2(n_s + n_m)$ is

---

8 Efficiency also vanishes under a different assignment of property rights. In particular, if the midcity bridge were controlled by residents of the central zone, and if the suburban bridge were controlled by midcity residents, then the groups (not being users of the bridge they control) would be indifferent to bridge capacities. However, the more plausible property-rights assignment is the one assumed, under which bridge capacities are controlled by those with a stake in the outcome.
replaced by $T_2^L(n_s + n_m)$, given that the low-income voters who live in the midcity zone make the bridge’s capacity choice. As a result, the equilibrium and optimality conditions for this choice ((24) and (17)) are no longer the same, so that bridge capacity is chosen inefficiently. Assuming that $T_2^H < T_2^L < 0$ holds, so that additional capacity reduces time-cost by less for the low-income group, the bridge capacity will be set too low, conditional on the equilibrium population distribution. The reason is that, although cost sharing partly internalizes the capacity spillover, it fails to capture the suburban residents’ greater benefit from additional capacity, with capacity instead chosen solely on the basis of low-income time costs. This failure of efficiency identifies a limitation of cost sharing as a remedy for spillovers: it only works when the preferences of the decision makers are the same as those of the individuals benefiting from the spillover.9

IV. CENTRALIZED DECISIONS

While decentralized capacity choices are efficient under the conditions of Proposition 1, it is illuminating to consider the centralized case, where a single city government sets capacities and fees on both bridges, under these same conditions. Before characterizing choices in this case, let $T(n_s + n_m, k_m)$ be written as $t[(n_s + n_m)/k_m]$ using the zero-degree homogeneity assumption, where $t'>0$ ($T^S$ similarly becomes $t(n_s/k_j)$). Then, (16) and (17) reduce to

$$
\frac{\beta k_m}{n_s} = t' \left( \frac{n_s}{k_s} \right) \frac{n_s + n_m}{n_s + n_m} = t' \left( \frac{n_s + n_m}{k_m} \right) \frac{n_s}{k_s}.
$$

Given the common form of these conditions, it follows that the volume/capacity ratios $(n_s + n_m)/k_m \equiv R_m$ and $n_s/k_s \equiv R_s$ on the midcity and suburban bridges are the same in the decentralized equilibrium, taking common a value denoted $R^*$. Congestion levels and user fees are then also equalized.

The equal-congestion condition could be imposed as a horizontal-equity requirement in centralized decisions, with volume/capacity ratios on both bridges required to equal some common value $R$. If equal user fees were also required, ruling out cross-subsidies between bridges, the fees would equal $\beta R$ on both bridges. Transport costs would then be $(\beta/R) + t(R)$ for midcity residents and twice this value for suburban residents, leading both groups to prefer a common $R$ equal to $R^*$. With central residents indifferent to $R$’s value, all the city’s residents would then unanimously support choice of $R^*$, implying that the efficient, decentralized outcome is achieved under centralization. Summarizing yields

**Proposition 2.** If horizontal equity is required, with both congestion levels and user fees constrained to be equal across bridges, then centralized choice replicates the efficient decentralized equilibrium.

9 While the optimum and the equilibrium with income heterogeneity are not compared, note that the zonal homogeneity assumed to characterize the equilibrium will generally not be preserved, with the $H$ and $L$ groups most likely mixed in either the midcity or suburban zones.
By contrast, suppose that a uniform user fee $\tau$ that balances the central budget were required under centralization, but that bridge congestion levels were allowed to differ. The user fee would then satisfy $(2n_s + n_m)\tau = B(k_s + k_m)$. Midcity and suburban total user-fee payments would then be $B(k_s + k_m)/(2n_s + n_m)$ and twice this amount, respectively. Substituting these values into (12) and (13) demonstrates that residents of two zones would have different preferred values of $k_s$ and $k_m$. These preferences would have to be aggregated in some fashion in a centralized decision process, whose outcome is certain to be inefficient.

V. CONCLUSION

This paper has shown that the inefficiency of fiscal decentralization in the presence of spillovers, a main tenet of the decentralization literature, is overturned in a particular transportation context. In a monocentric city where road (bridge) capacity is financed by budget-balancing user fees, decentralized capacity choices (made by individual zones within the city) generate the social optimum despite the presence of spillovers. Optimality also requires the correct population distribution across the city’s zones conditional on bridge capacities, and this outcome is achieved because the user fees function as optimal congestion tolls, a result that follows from the famous self-financing theorem of transportation economics. Therefore, in addition to showing that efficient decentralization can occur with spillovers, the paper shows for the first time how optimal tolls can emerge in a decentralized setting in a monocentric city. This emergence, however, is crucially tied (via the self-financing theorem) to the simultaneous optimization of bridge capacities.

While the paper’s efficiency result can be overturned when the technological and user-fee assumptions are relaxed, the fact that these assumptions are fairly realistic may reduce concerns about this result. Of greater concern is the loss of efficiency that arises from income heterogeneity, which is a hallmark of actual cities. This loss reflects a general inability of user fees to fully remedy the inefficiency due to spillovers when decision makers and the individuals experiencing the spillovers have different preferences. While cost sharing via user fees helps with the spillover problem, the reliance of decision makers on their own preferences, ignoring the difference in the valuations of outsiders, prevents achievement of a fully efficient outcome. This issue, rather than the possible inaccuracy the model’s specific assumptions, may constitute the main barrier to the broad relevance of its lessons.

Nevertheless, the analysis highlights the role of user fees as at least a partial remedy for underprovision of goods and services that generate beneficial spillovers. With decentralized decisions, the cost sharing that user fees entail spurs an expansion of output relative to the case of own-financing, and even though this expansion may be insufficient or excessive due to decision makers’ failure to account for differing preferences, it may close much of the gap between the optimal and equilibrium levels of provision. Moreover, the paper shows that budget-balancing user fees work perfectly when preferences are homogeneous, and they function as optimal tolls when the service they finance is congestible.
ACKNOWLEDGMENTS

I thank Jinwon Kim, Kangoh Lee, David Merriman, Ken Small, Yukihiro Kidikoro, David Wildasin, several referees, audience members at the 2014 ITEA conference, and especially Sofia Franco for comments and suggestions. Any errors, however, are mine.

DISCLOSURES

The author has no financial arrangements that might give rise to conflicts of interest with respect to the research reported in this paper.

REFERENCES


