A MARKET-ORIENTED TRANSPORTATION AND LAND USE SYSTEM:
HOW DIFFERENT WOULD IT BE?

Douglass B. Lee¹
US Department of Transportation,
USA

ABSTRACT

If transportation operated according to market principles, how different would the world be? In particular, would the types of transportation provided and the spatial organization of cities be much different? The proposal is to explore the possible ways in which transportation infrastructure and land use patterns would change under a set of ideal conditions or "rules" applying to both transportation markets and land use markets. The rules are those by which a "pure" market system would operate.

OUTLINE OF THE ARGUMENT

The primary focus is on urban or metropolitan land use, and hence primarily passenger transportation, although the argument extends into freight and intercity markets.

The Hypothesis. The hypothesis is that efficient transportation and land use markets would yield substantially different results -- most notably, higher vehicle occupancy, less highway mileage, a higher transit mode share, and more compact urban development -- compared to what currently exists. In many ways, it would seem, the objections to existing transportation and land use patterns voiced by various groups -- too much congestion, too many highways, wasteful transit subsidies, poor transit service, high energy consumption, environmental degradation, wasteful land use patterns -- would, despite the apparent positions of the groups on opposite sides of a wide range of issues, be at least partially resolved.

This hypothesis stands in partial opposition to both of the two dominant schools of thought regarding urban transportation policy. On one side are those who assert that urban form should be shaped by public policy, including transit subsidy; on the other side are those who assert that such policies have little effect and waste resources. The first school seems to imply that the market is the enemy, and should be overruled.

The second school seems to suggest that the market works just fine and should be left alone.² The hypothesis presented here supports some of the objections to existing transportation and land use patterns voiced by the first school, but agrees with the second school in deducing that present policies are not the way to achieve such worthwhile goals. Markets work well when given the right information, but existing urban transportation and land use markets are distorted by market failure and government interventions described below.

¹ U.S. Department of Transportation, Transportation Systems Center, Kendall Square, Cambridge, MA 02142. Previous versions of this paper have been presented at meetings of the Northeast Regional Science Association at Cornell University, October 1988, and the Transportation Research Board in January 1989.

² It is a bit unfair to caricature works of literature according to positions that the authors never had in mind, but the first school is represented by planners, environmentalists, and transit advocates, while the second tends to be economists, engineers, and political scientists. Some examples that tend toward the first school are Audirac et al. (1990), Brown and Jacobson (1987), Cushman (1988), Lowe (1991, 1992), MacKenzie et al. (1992), Newman and Kenworthy (1989), Register (1987), Renner (1988), Wingo (1961), and Yaro (1991). Some that tend toward the second school are Boyce (1980), Domencich and Kraft (1970), Gomez-Ibanez (1985), Gordon and Richardson (1989), Ingram (1979), Kain (1983), Knight and Trygg (1977), Love and Cox (1991), Meyer and Gomez-Ibanez (1981), Netzer (1970), and Small (1985). The former group is in favor of making autos pay their "full cost," but not necessarily willing to accept the outcome of "fair" markets. The second group has been primarily concerned with attacking myths about transit and land use -- transit saves energy and reduces air pollution, better transit service will attract riders from autos, transit is more "efficient" in carrying passengers, land use plans and transit investment can shape urban form -- but supports the principles of efficient pricing and the benefit-cost criterion for investment. One of the more balanced views is offered by Pucher (1988).
The Argument. Technically, the standard being applied here is referred to as efficiency. All resource decisions would be made through efficient markets, rather than through political processes. The claim is that a market-based transportation and land use equilibrium is both desirable in its characteristics and different from what we have now. Therefore the argument has three components:

(1) **Definition of Efficiency.** The rules by which the system would operate require that price equal marginal cost in all relevant markets, that all investments in capacity be able to pass a benefit-cost test, and that all transportation or land use enterprises support themselves entirely out of their own revenues.

(2) **Inefficiency of Existing Prices and Investment.** Existing U.S. transportation and land use markets are distorted as a consequence of uncontrolled externalities (market failure), overinvestment in and underpricing of public infrastructure and services (government failure), and other malfunctions. State, local, and national policies with respect to land use, zoning, housing, taxation, regulation, and subsidy tend to reinforce the inefficient supply and overconsumption of transportation services and land.

(3) **Magnitude of the Behavioral Response.** Demand elasticities are such that correct pricing and market-constrained investment would result in much different transportation consumption, transportation investment, and land use patterns. The changes would be large because existing distortions are consistently in the same direction; for the most part, the analytic problem is not one of evaluating how strong are some effects in relation to countervailing influences, but how large is the sum of all such effects.

Given the large bundle of policies and programs that would be altered under efficient pricing and investment, and the extensive indirect effects such changes would have, it is impossible (and would be misleading) to attempt to offer a rigorous derivation of predicted impacts. The argument is one which depends upon a systematic review of theory and empirical evidence, and a judgmental synthesis of their quantitative magnitudes.

**EFFICIENCY: CRITERIA AND ASSUMPTIONS**

(1) **Marginal Cost Pricing (Utilization).** Transportation output is priced at marginal cost. The price of usage, per vehicle or person trip, is based on direct variable costs plus imposed delay, which yields the price that rations available capacity to the optimum flow,

\[ P_i = MC_i \]

where \( i \) represents each relevant transportation market. An analogous short-run criteria applies to land use transactions such as public services. This rule implies that whatever capital stock is available -- at present or in the future -- is optimally utilized.

(2) **Benefit-Cost Investment Criterion (Capacity).** Capacity is determined so as to maximize incremental minus costs, considering only the impacts in the direct market served:

\[ B_i > C_i \]

where \( j \) denotes each individual project, ordered in a way so as to properly account for interaction effects among projects. This criterion applies to long run (i.e., investment) decisions regarding both transportation and land use facilities.

(3) **Full Cost Pricing.** This is a constraint imposed on efficiency for institutional reasons: the motivation to control costs and make wise investment decisions is believed to be greater in the long run if each enterprise must survive out of its own revenues, so not transportation mode or sector is directly or indirectly subsidized. Each mode, each functional class of highway, etc., is supported entirely from user charges and other operating revenues. Multi-part prices are used to recover residual costs from users without distorting their consumption, if marginal cost prices fail to produce sufficient revenues. Economically, every tub stands on its own bottom, so that

\[ TR_k = TC_k \]

which says that total revenues for enterprise \( k \) must equal total costs for the same enterprise. Costs are full long run costs, including general taxes paid to each level of government, and revenues are dollar revenues. An enterprise might be a road authority, a transit operator, a land use, or a household. It is more important that enterprises be self-supporting, and make their own independent business decisions, than that prices conform exactly to theoretically pure marginal cost pricing principles.

---

1 This means that the demand curve for the facility in question fully reflects all changes that occur in related markets. Thus benefits measured on the one facility take into account congestion on parallel and complementary facilities, land use changes nearby and elsewhere, and other related effects.
(4) **Optimal Externalities.** All negative externalities are optimally controlled, preferably by internalizing costs and benefits or alternatively by efficient regulation that minimizing control-plus-damage costs. Land use controls and environmental regulations, to the extent they are imposed, have no resource allocation effects beyond the control of externalities.

(5) **Other Assumptions.** Other market imperfections not elsewhere covered (such as scale economies, natural monopoly, nonrival consumption, merit goods, transaction costs, etc.) are assumed to be quantitatively minor and are adequately controlled or compensated for. All resources are mobile, meaning that any investment or disinvestment response lag is short enough to permit convergence on efficiency. Government policy is neutral with respect to choice of mode or land use. All benefits of transportation are assumed to be captured by users, and either consumed directly or passed on to indirect consumers. There are no external benefits from transportation that fall outside normal market processes. Thus transportation investment and land use patterns are determined by individual consumers, through expression of their preferences for travel, housing, life style, and other goods and services, in normal markets.

**EVIDENCE OF INEFFICIENT PRICING AND INVESTMENT**

The only basis for trying to guess what the outcome of an efficient process of transportation and land use decisionmaking would be is to extrapolate from existing conditions. The sources of distortions, the markets they occur in, and the factors that are distorted -- listed in Table 1 -- are somewhat complex in their interrelationships. Those particular categories that are helpful in describing a given distortion depend upon the nature and location of the distortion, so only those relevant to the particular context are explicitly mentioned in the discussion below. The general approach is to structure the problem in descending order of influence, so dominant factors are addressed first. Thus highways are central to the outcome, and passenger travel is more critical than freight.

<table>
<thead>
<tr>
<th>Table 1: A Typology of Existing Market Distortions</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) <strong>Market Failure.</strong> Problems with the proper functioning of markets that are inherent or technological, such as monopoly, scale economies, positive or negative externalities, and public goods, are called market failure. The list of such problems is not mutually exclusive, because the same example of failure can often be described in terms of different conceptual categories of failure.</td>
</tr>
<tr>
<td>(2) <strong>Government Failure.</strong> Government enterprises may underprice and over-invest without suffering the consequences of going out of business, and resource allocation decisions are often politically driven. Governments can introduce other distortions through regulation and taxation.</td>
</tr>
<tr>
<td>(3) <strong>Transportation Mode.</strong> Modal markets are divided into private auto and transit, consistent with the emphasis here on urban passenger travel.</td>
</tr>
<tr>
<td>(4) <strong>Supply and Demand.</strong> The level of transportation demand can be characterized as urban or rural, and congested or uncongested within urban areas.</td>
</tr>
<tr>
<td>(5) <strong>Short or Long Run.</strong> Distortions can be divided into those that affect pricing (short run utilization of existing capacity) and those affecting investment (construction or replacement of capital facilities).</td>
</tr>
<tr>
<td>(6) <strong>Land Use.</strong> Markets determining land use can be classified as those that shape urban form (macro influences) and those that address neighborhood externalities (micro-spatial effects).</td>
</tr>
</tbody>
</table>

**Congested Urban Highways.** Highways produce two fundamental kinds of output. One is space for vehicles with which to move people and goods; the other is strength to carry heavy loads. Very few passenger vehicles place much weight stress on the highway, and their demand is for space. Congestion is due to excessive consumption of space relative to the amount available. Rapid pavement deterioration is the consequence of axle
loadings much heavier than the strength of the road. In urban transportation, the problems concern space capacity, or volume capacity, more than weight capacity.

Peak travel in the peak direction, in urban and suburban areas, constitute the times and locations where transportation markets show evidence of underpricing. Underpricing that leads to congestion not only wastes time and other resources, it handicaps the productivity of bus or van transit services, both because they cannot offer competitive service (since the vehicles suffer the same delay) and because the vehicles are able to make fewer trips during the peak period. Despite the crowding of vehicles on the road, vehicle occupancy is absolutely low (about 1.3 persons), leaving generous amounts of space for additional passengers.

**Uncongested Highways.** Travel at non-peak times and in the non-peak direction during peak periods does not suffer much congestion, even on urban facilities. Rural facilities are effectively uncongested at all times. With average user charges and user taxes (most of which are fixed with respect to short run travel and none of which correlate with congestion) amounting to about 1.7 cents per VMT, any lack of revealed demand cannot be due to excessively high user fees.

In fact, the level of utilization of most U.S. highway mileage is very low. Over half the mileage carries less than two hundred vehicles per day, and significant shares of all functional classes are underutilized when the long run cost of the highway is taken into account as well as traffic volumes. The average traffic volume on the entire Rural Collector and Local system (75% of the US highway mileage) is less than 300 vehicles per day, a flow that could be accommodated in ten minutes of full capacity operation on these roads.\(^4\) for the most heavily utilized category, urban primaries, the full daily load is the equivalent of 4.5 hours at capacity flow. Over one-third of the Rural Interstate mileage carries less than 10,000 vehicles per day -- an average density of two or three vehicles per lane mile -- which is a volume that can readily be carried on a two-lane road. Removing the least-productive 75% of highway mileage would affect (i.e., cause a detour for) barely 13% of the traffic, yet reduce capital expenditures by over half. Most highways are not congested, and even congested ones are uncongested most of the time.

Table 2 shows several other indicators of utilization. Capital costs are based on a single prototype for each functional class, as a means for estimating the average replacement cost. When amortization and “interest” are incorporated, the long run average cost per VMT ranges from five cents on Interstates to forty cents for the low-volume rural local service system.\(^5\) Although the lowest of these unit costs is about four times what users currently pay, a well-utilized highway could cover its full capital (and presumably its operating) costs out of user revenues. The unit cost problem is too few vehicles, not costly roads.

\(^4\) These ratios are derived from Highway Statistics 1990, and have been calculated previously, using 1987 data, in Lee (1989). The rural collector and local system contains about 67% of the physical capacity in the highway system, but less than 30% of the capital value.

\(^5\) If highways are viewed as economic investments, analogous to the way a business views an investment, then there are opportunity costs that are largely or entirely ignored in the published accounts of highway expenditures. Two of the most important of these are land opportunity costs and borrowing costs. Land in highway right-of-way has alternative uses, and this value is included in published figures only when the purchase of new land is a part of current expenditures. Normally, any long-lived business investment is expected to earn a rate of return at least equal to the interest rate on borrowed funds; with pay-as-you-go accounting, interest costs are largely omitted.
Table 2: Excess U.S. Highway Mileage by Functional Class

<table>
<thead>
<tr>
<th>Functional Class</th>
<th>1990 Capital Average ADT</th>
<th>Capital Cost ($000/mi)</th>
<th>Long Run Cost ($/VMT)</th>
<th>$.20 Threshold Volume</th>
<th>% Excess Mileage</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>INTERSTATE</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>16,380</td>
<td>3,836</td>
<td>0.06</td>
<td>4,599</td>
<td>10</td>
</tr>
<tr>
<td>Urban</td>
<td>66,171</td>
<td>15,257</td>
<td>0.05</td>
<td>18,019</td>
<td>12</td>
</tr>
<tr>
<td><em>MAJOR ARTERIAL</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>5,734</td>
<td>1,900</td>
<td>0.08</td>
<td>2,169</td>
<td>25</td>
</tr>
<tr>
<td>Urban</td>
<td>21,269</td>
<td>10,245</td>
<td>0.11</td>
<td>7,936</td>
<td>19</td>
</tr>
<tr>
<td><em>MINORAL ARTERIAL</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>2,950</td>
<td>1,216</td>
<td>0.10</td>
<td>1,499</td>
<td>36</td>
</tr>
<tr>
<td>Urban</td>
<td>8,625</td>
<td>4,355</td>
<td>0.12</td>
<td>5,187</td>
<td>39</td>
</tr>
<tr>
<td><em>MAJOR COLLECTOR</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>1,201</td>
<td>780</td>
<td>0.16</td>
<td>978</td>
<td>64</td>
</tr>
<tr>
<td>Urban</td>
<td>3,633</td>
<td>4,367</td>
<td>0.29</td>
<td>3,161</td>
<td>58</td>
</tr>
<tr>
<td><em>MINOR COLLECTOR</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>470</td>
<td>520</td>
<td>0.28</td>
<td>661</td>
<td>81</td>
</tr>
<tr>
<td><em>LOCAL SERVICE</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>125</td>
<td>201</td>
<td>0.40</td>
<td>250</td>
<td>86</td>
</tr>
<tr>
<td>Urban</td>
<td>1,011</td>
<td>774</td>
<td>0.18</td>
<td>913</td>
<td>72</td>
</tr>
</tbody>
</table>

ADT = annual vehicle miles of travel (VMT) divided by highway mileage, for each functional class, from Highway Statistics 1990.

Capital Cost = FHWA (HPMS) estimates of right-of-way (ROW) cost per highway mile, for a representative cross-section, plus pavement reconstruction cost per mile, both for 1978 updated to 1990 using the FHWA Highway Construction Composite cost index.

Long Run Cost = estimated annual capital cost of the functional class, using capital recovery factors based on a 7% discount rate with a lifetime of 20 years for pavement construction and infinite life for ROW, divided by annual VMT for the functional class.

Threshold Volume = ADT for the functional class below which the long run cost is higher than $.20 per VMT.

Excess Mileage = percent of highway miles in the functional class falling below the threshold volume.
The “threshold” volume is the result of applying an arbitrary standard of $.20 per vehicle mile to the volume distribution for each functional class. This is a level of user fee that would be high enough to ration flow on all but the most highly congested urban facilities, and would severely reduce traffic volumes on any functional system to which it was applied. Urban systems probably could not meet this average revenue even with fairly high congestion charges, and rural systems would need to reduce average cost by consolidating mileage. Thus a threshold of twenty cents per VMT is generous with respect to the retention of existing mileage. Using this test, however, indicates a surplus of about 10% in the Interstates and over 80% for rural secondary classes, with other functional classes ranged in between.

Figure 1 shows ADT distributions for the two ends of the functional class spectrum, all of which have the bulk of their mileage at the relatively low ends of their potential capacity.

**Figure 1**
Parking Pricing, Parking Capacity, and Zoning Requirements. Similarly to highway infrastructure, parking is overprovided and underpriced, particularly in those markets where transit and carpooling are the most feasible. Cities have built parking lots and garages in the apparent belief that free or subsidized parking is a public service that benefits citizens and the local economy. Even when cost are ostensibly recovered from users, the costs of land and borrowing are understated. In part because the imputed value of parking is not taxed, employers provide parking as an employee perquisite. Similarly, unions demand free parking as a fringe benefit. Property tax practices that tax actual improvements rather than the potential for intensive use (i.e., a land tax) suppress development below its optimal intensity, thus favoring parking as a land use.

Imposed through zoning regulations, new development has been forced to incur costs for parking provision which have not been recovered from parking users. Standards which require a minimum number of parking spaces per resident or employee have gradually given way in some urban areas to those imposing a maximum.

Environmental Externalities. Air pollution, water pollution, noise, danger to pedestrians, and visual degradation are some of the many broad categories of negative impacts caused or contributed to by highways that are unpriced to users. Some are variable (emissions leading to smog, road salts), and some are fixed (impervious surface, wetland destruction) with respect to usage. Estimates of these damages have placed the value in the two-to-three-cents-per-VMT range, but the estimates are conservative, incomplete, and generally not based on recent data. Many people believe that highways exert subtle but negative influences on community and lifestyle values, although it is hard to know to what degree highways are a causal factor and what positive effects are perhaps being overlooked.

When attention is focused on selected aspects, however, it seems clear that uncontrolled highway externalities correspondingly degrade related activities. For examples, pedestrians are frequently accommodated within close physical proximity to streets and highways, and the greater the noise and other impacts from traffic, the less attractive is the pedestrian environment. The institution of traffic “calming” measures reflects an attempt at restoring a balance of neutral treatment for walking as a mode of transportation.

Property rights for externalities are poorly defined, legally, and protection has depended upon public regulation and enforcement, rather than trading in economic markets. The result is undercontrol (despite overregulation) and inefficient control (increased production costs that have limited effect on externalities). There are no incentives to decouple the externality from the associated activity, so regulation aims at reducing output (such as VMT) at least as much as at reducing the externality (emissions).

Declining Transit Demand. The history of transit is one of declining underlying demand since the 1920s, inversely proportional to the growth of the highways and use of the automobile for passenger travel. The period of the Second World War was an artificially high level of transit use, due to severely restricted supplies of fuel and cars. The more gradual secular decline experienced since the 1950s has been masked by variations in subsidy and fare policies, with ridership occasionally increasing slightly as a consequence of fare reductions and service expansion, enabled by increased subsidies.

Three broad causes can be assigned responsibility for the decline in transit demand. First, there is an inherent preference for auto travel facilitated by increased real incomes and reduced real cost of auto ownership and use. Second, transit is produced and deployed by public agencies, with the poor market orientation that political control entails. Third, there are numerous land use policies which favor highways and auto travel over transit travel, in addition to low-price highways. Each of these causes has contributed, independently, to transit’s decline, although the proportionate shares are difficult to measure.

It is inevitable that the auto would capture market share from transit, as auto costs came within reach of most families. Even had transportation policies been completely neutral with respect to choice of mode and transportation consumption in relation to other goods, the auto would have gained market share. As a factor in the decline of transit, however, secular decline in demand is perhaps the least of the three.

---

6 Pioneers in asserting the influence of parking pricing on mode choice are Pickrell and Shoup (1980), Pickrell (1990), and Shoup and Willson (1992). They report, for example, that 74% of federal employees who park in downtown Washington pay nothing, and 97% of the firms submitting air quality plans in Southern California offered free parking to all employees.

7 The literature up to 1982 is reviewed in U.S. FHWA (1982), and a more recent survey is Small (1991).

8 Current regulation of vehicle emissions, for example, places more stringent requirements on new vehicles, which are already relatively clean, and grandfathers older vehicles. If it is true that 10% of the vehicle fleet produces 50% of the pollution (the Stedman hypothesis), then costs are incurred by many vehicle owners for which there is insignificant improvement in air quality.
Transit Production Inefficiency. Operators have been constrained from exercising effective cost management, from creating sensible fare structures, and from optimizing service design and deployment. Even when not constrained, incentives for good management are weak and performance of the industry has been mixed at best.\textsuperscript{9} If transit operators were to achieve competitive business levels of labor compensation, productivity, market responsiveness (service quality and reliability, pricing), and cost control, the transit industry would be able to provide better service at a lower cost. Whether ridership would be much higher is harder to assert, but it seems clear that more could be done with the available resources if they were focused primarily on delivering transportation where it is demanded.

Transit's market is most heavily impacted by the underpricing and overbuilding of highways, particularly since the greatest potential for transit is in exactly those corridors where highway capacity is underpriced to the greatest degree. Attempts to restore balance by direct and indirect subsidy, rather than by increasing the price of commuting and parking, only make the distortions worse. Market conditions for transit will be much more favorable under full pricing for both highways and transit than with both modes subsidized.

Government Land Use Policies and Service Provision. Compared to historical patterns, present urban areas are multi-nodal and spread out. Seemingly underutilized land is common, as well as destructively incompatible land uses. If land is being treated as less costly than other resources, such that firms and households choose to acquire new land, perhaps somewhat removed, rather than fix the problems on a given parcel, then it is likely both that transportation is cheap and that neighborhood externalities are poorly controlled.\textsuperscript{10}

Subsidies to dispersal, such as mortgage interest deduction, subsidized interest rates for home ownership, and government absorption of bad loans have also tended to favor auto travel. Infrastructure other than transportation (sewers, water, fire and police protection), and services (education, solid waste), tend to be overextended due to government subsidies, and underpriced. Even when revenues cover full costs (rare in government), pricing is flat with respect to location, overcharging easily serviced customers at the expense of remote ones. Zoning serves to obstruct land use clustering by artificially segregating land uses, forcing lower density and greater distances between trip ends for the same density.

Neighborhood Land Use Externalities. Attempts to reduce the nuisance effects of some land use activities on others have been directed into the institutional mechanism of zoning, which has diverted any effort to resolve such incompatibilities through property rights and private legal redress. The result is that negative neighborhood externalities have been tolerated when they shouldn't be, because zoning grandfathers pre-existing situations and because political enforcement is weak.

Slow-Growth Movements. Again using the institution of zoning, and state-level land use statutes such as growth boundaries and planning requirements, existing residents exercise property rights over developing property with the effect of suppressing density and forcing development further to the fringes of metropolitan areas.

OUTLINE OF AN EFFICIENT EQUILIBRIUM

Several generally accepted theories will be briefly presented as a basis for judging the net effects of changes in transportation pricing and investment. These theories concern congestion pricing, efficient investment in capacity, and land use structure.

Highway Congestion Pricing Theory. In order to conclude that highway travel would be priced at substantially higher levels under both efficient pricing and efficient investment requires several logical steps. First, congestion implies the price is too low. Second, investment in capacity lowers the efficient price, no matter how elastic is demand. Third, the efficient price might be lower than the existing price after efficient investment, if the cost of additional capacity is low enough.

Figure 2 represents both the efficient pricing of a given facility and the incremental benefits of expanding capacity from a smaller to a larger facility. Marginal costs are based only on variable costs, mainly travel time, running costs, and accidents. Congested highways are usually operated at a point close to the intersection of the average cost curve (e.g., $AVC_o$) and the demand curve. Clearly, the efficient price for the initial facility is $p_o$, which will always be lower than the existing price without tolls. Since the existing cost to the user includes delay time, the efficient toll ($p_o-c_o$) substitutes a money price that is higher than the average value of the delay cost.


\textsuperscript{10} Neighborhood externalities are negative effects such as noise, smoke, dust, and visual degradation that occur "in the neighborhood of" the source, i.e., as a result of close physical proximity.

226
Previous estimates place the correct price at around ten cents to a dollar per vehicle mile in today's dollars, depending upon the level and elasticity of demand. The only comprehensive national study places the average efficient congestion charge on urban Interstates at about ten to fifteen cents per vehicle mile, and estimates that efficient pricing would raise $75 billion annually from congestion charges alone. Current user taxes raise slightly over $68 billion from all sources and all levels of government, and none of these are congestion-oriented.11

**Investment in Capacity.** On the larger facility described by $MC_1$ and $AVC_1$, the efficient price is lower than the inefficient actual price on the smaller facility. Do situations exist where capacity expansion is warranted that will reduce the efficient price? Unit costs (per trip, or per mile) at free speed volumes are reduced from $y_0$ to $y_1$ with the expansion, due to higher speeds, lower running costs, and greater safety. Delay savings on old trips measured from an efficiently priced base are generally modest. As demand grows, the value of the expansion is in its ability to serve new trips. The efficiency of the investment depends upon whether the shaded areas12 -- summed over all demand periods for all times of day for the life of the facility -- exceed the incremental costs. If yes, the price on the expanded facility could be less than the current inefficient price.

---


12 Incremental net operating benefits (i.e., excluding capital or fixed costs) are represented by the area between the two marginal cost curves and also underneath the demand curve, or $y_0p_o y_1$ in the diagram. Because of the relationship between the marginal and the average cost curves, this area is equivalent to the shaded area.
Table 3 shows the results of a prototype project evaluation, in which a set of parameters were designed so as to create the minimum benefits required to justify investment in capacity expansion at a given cost. The highway type is nominally a four-lane expressway. If a highway can be built for $5 million per mile (first column of numbers) that will carry 16,717 vehicles per day at a maximum volume-to-capacity ratio of at least .60, then it is worth doing. Costs and benefits are assumed to be incremental measured from the base case of an existing two-lane arterial. Annual hours of demand are divided into peak, off-peak, and residual, so as to yield K-factors (the ratio of peak hour to ADT volume) similar to existing highways. Thus the $5- and $10-million facilities are rural, and the others urban. If the capital cost values in Table 2 are correct, the results of the prototype analysis indicate that the average ADTs for Interstate highways are sufficient to justify their construction, but the mileage labelled "Excess" falls well below a benefit-cost threshold. The conclusion is that a few additional urban expressways might possibly be feasible, but they will require substantial volumes of traffic willing to pay congestion tolls of over $.25 per VMT. Because only capital costs are included, when other fixed and variable costs are added in, the elasticity of demand may be such that the amount of mileage eliminated by efficient pricing and investment would greatly exceed the amount added, even in congested urban corridors.

Table 3:
Prototype Benefit-Cost Evaluation of Highway Capacity Expansion

<table>
<thead>
<tr>
<th>Capital Cost ($mil)</th>
<th>5</th>
<th>10</th>
<th>20</th>
<th>50</th>
<th>100</th>
<th>100*</th>
</tr>
</thead>
<tbody>
<tr>
<td>AADT 4-Lane</td>
<td>16,717</td>
<td>31,466</td>
<td>40,437</td>
<td>52,295</td>
<td>80,682</td>
<td>112,937</td>
</tr>
<tr>
<td>AADT 2-Lane</td>
<td>10,530</td>
<td>21,641</td>
<td>28,703</td>
<td>39,482</td>
<td>57,391</td>
<td>55,346</td>
</tr>
<tr>
<td>K-factor</td>
<td>0.14</td>
<td>0.12</td>
<td>0.10</td>
<td>0.08</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>Peak V/C</td>
<td>0.60</td>
<td>0.95</td>
<td>1.03</td>
<td>1.04</td>
<td>1.04</td>
<td>1.02</td>
</tr>
<tr>
<td>Off-Peak V/C</td>
<td>0.36</td>
<td>0.57</td>
<td>0.62</td>
<td>0.62</td>
<td>0.83</td>
<td>0.82</td>
</tr>
<tr>
<td>Peak Demand Elasticity</td>
<td>-1.00</td>
<td>-0.80</td>
<td>-0.70</td>
<td>-0.60</td>
<td>-0.60</td>
<td>-0.60</td>
</tr>
<tr>
<td>Annual Peak Hours</td>
<td>200</td>
<td>500</td>
<td>1,000</td>
<td>1,850</td>
<td>2,750</td>
<td>2,250</td>
</tr>
<tr>
<td>Annual Off-Peak Hours</td>
<td>800</td>
<td>1,500</td>
<td>2,000</td>
<td>3,000</td>
<td>4,000</td>
<td>3,500</td>
</tr>
<tr>
<td>Efficient Peak Toll</td>
<td>0</td>
<td>0</td>
<td>0.26</td>
<td>0.65</td>
<td>0.95</td>
<td>0.38</td>
</tr>
</tbody>
</table>

4-Lane Capacity = 4,000 vph
2-Lane Capacity = 2,000 vph
Value of Time = $6.00/hr
Marginal Cost excluding Time for 4-Lane = .23/vmt
Marginal Cost excluding Time for 2-Lane = .33/vmt
Capital Recovery Factor = .10

Prices After Rationalizing Capacity. An expanded facility would reduce the price to the user and add more trips. Consolidating excess highway mileage on low-volume networks into a reduced but more fully utilized system would reduce long run costs. The problem, then, is to guess what levels efficient highway prices would take if the investment were fully rationalized.

With the investment evaluation, the three pieces of the puzzle are in place. It is clear that average user charges need to be much higher in percentage terms than they are now, and that large portions of the total system -- especially in the lower rural functional classes would be deleted from an efficiently invested system.

Prices per vehicle mile would be in the $.10-.20 range for high-demand urban and suburban roads, and in the $.05 range on a highly consolidated rural network. Total mileage will be much less on an efficient network.

Demand Elasticities. If highway prices were higher, would people behave much differently, especially since the level of service would be generally higher? There is plenty of evidence that highway demand is elastic,
even when transit is not an option. Parking pricing alone, for example, would cause significant alterations in mode choice and commuting patterns.

The key to understanding demand is to realize that efficient pricing effectively never places a charge on the person-trip itself, but rather, on some resource indirectly consumed. For highways, the resource is the space on the highway and the externalities produced by the vehicle and the highway, and the charge falls on the vehicle. Thus travelers can escape the charge by sharing vehicles, or by shifting to a less-demanded corridor or time of day. Whether users choose to carpool or ride transit, vehicle occupancy changes could absorb most of the vehicle-price increase without seriously affecting person-trip demand. Vehicle price elasticity is probably close to -1.0.

Land Use Theory. Many urban theories incorporate aspects of location, transportation, land use, and activity types. Most of these are microscopic, in that they deal with particular locations and prototypical activities. The most useful theory for relating transportation and land use is highly aggregated, i.e., it is a macro theory. It is derived from rent theory, and it yields the urban intensity gradient.

A macro model is intended to explain a property of the whole that is not simply a consequence of summing up individual elements. Such a theory does not substitute for the explanation of localized variations or patterns addressed by microscopic models. The faults of the urban rent model are not -- as is often claimed -- that it is unrealistic (in being "monocentric") or inaccurate (in explaining microscopic land values or densities), but that it has not been sufficiently abstracted so as to explain, for example, land use "intensity."

In the model outlined here, the "market" at the center of the urban area generates the demand curve for the output of land-intensive activities. The market and the rent gradient are shown in Figure 3. The supply curve is the area of a circle whose radius is the market price divided by the transport rate, that being the outer extent of the production area. A shift in the demand curve outward will increase the market price and raise the rent gradient, at the same slope. A reduction in the cost of transportation, in contrast, will flatten the rent gradient. Assuming the demand schedule does not shift, total output will increase, land near the center will decline in value, and land at the fringe will increase in value. More total land will be in production.

Figure 3

13 High fuel taxes in Europe are sometimes cited as evidence that demand is inelastic. A tax of $2 per gallon is about 5 cents per mile after fuel mileage adjustments take place. This magnitude undoubtedly has an effect on travel, but it is not a high price, and it is not a substitute for efficient user charges. In practice, it is an excise tax on a particular commodity and is largely used for general government revenues. As a general tax, it is probably not an excessive rate to charge highway users, in addition to congestion and pavement damage charges.

14 One of the earliest contributions to the transit-auto cross-elasticity debate was Domenich and Kraft (1972), which concluded that it would be necessary to pay people to ride transit in order to lure them from their automobiles. The complete story obviously is more complex than this, but the basic truth is that auto travel is a small enough expenditure relative to most household incomes that transit cannot be made sufficiently vice. It is both feasible and defensible, however, to raise the price of highway travel to the point that demand is elastic and the cross-elasticities significant, which is what is proposed here.

15 The model is an extended von Thunen rent model, using features adapted from Alonso, Mills, and Muth.
The same flattening of the rent gradient will occur if the price of transportation to the user is reduced, even if the true social cost is unchanged. The subsidy must be extracted from somewhere, but the “where” is outside this model. The inefficiency of the underpricing, however, can be identified as the shaded area in the diagram. It is composed of production that is induced by expanding the urban area, and the transportation costs to bring it to the market, neither of which are justified by the benefits generated (the area under the demand curve).

**Land Use Intensity.** Rather than make distinctions among types of land use (residential, commercial) and calculating their densities (population, employment) as surrogates for land use intensity, the attribute “intensity” can be generated in the abstract. Land is simply one input to a production function, and the ratio of all other inputs to land can be called intensity. Other inputs include capital (to build taller buildings), labor (employment at a site), and materials, the last two categories requiring transportation and hence affecting location.

Using a simple Cobb-Douglas production function with two inputs and constant returns to scale, the intensity of land use is determined (simultaneously with the rent gradient) as a function of location. This yields the familiar concave rent gradient, although not specifically the exponential decay function commonly used to fit density gradients. The intensity-rent gradient is shown in Figure 4, for two transport price levels.

---

16 In this simplified model, production cost is zero, so the inefficiency consists solely of transportation costs for the excess production. The details of the basic rent model and the derivation of the intensity gradient are in Lee (1992).
**Agglomeration and Subcenters.** Contemporary cities do not have a single central location from which the rent/Intensity gradient declines smoothly with distance. Nor are all cities the same size. From the standpoint of rent theory, these effects are exogenous; the theory says that, given a demand for area-produced inputs at a specified location, rent and intensity will decline from there. Reasons why there might be big centers and small ones, or multiple centers in the same urban area, are found in agglomeration economies.

Activities would prefer, all other things being equal, that transportation be unnecessary. To accomplish this, all activities would have to locate at a single dimensionless (neither height nor breadth) point. The reason this cannot be is that activities take up space; they can be stacked vertically or arrayed horizontally, but they require transportation to provide physical interaction in any event. It is the physical interconnections -- movement of people and things -- that causes activities to want to locate close together.

Subcenters are local peaks in the intensity gradient. Each activity, whether residential or commercial, is evaluating tradeoffs with respect to what it wants to be near and how much it is willing to pay for that nearness. Subcenters permit activities to agglomerate at a smaller geographic scale -- the gradient is just as steep as in the larger centers, but it is lower and covers less area -- while still drawing upon access to the larger centers. Although these subcenter activities value high access, they value it relatively less than does the market as a whole, i.e., other activities. 17

A subcenter is thus defined by intensity, not land use or employment; although we do not observe residential subcenters, there is no reason in principle why residential activities might not want to cluster in an intense environment that provided multiple service and recreational amenities, along with remote high-speed access to employment and travel. Sub-centers may be central places (subregional), manufacturing nodes, or unique (e.g., Hollywood). To the extent there are macro explanations for the spatial structure of subcenters within urban areas, they lie in the concepts of Christaller and Losch. All rent and central place theories are constructed on an undifferentiated plain (or plane), so any specific locational attributes (topography, bodies of water, mineral deposits, microclimates) constitute distorting influences.

Cities were more easily described as monocentric prior to the introduction of the automobile. One hypothesis about the relationship of subcenters to the main “tent” is to posit that, so long as urban transport costs were high, potential subcenters were forced to bundle themselves together in a single central cluster. As transport costs to the user were lowered, costs were to be increased, the process would reverse: some centers would re-bundle themselves (not necessarily at the main center). 18

From the above discussion, the expected changes in the overall urban intensity gradient might look like those in Figure 5. Peaks in the tent would be sharper, perhaps closer together but fewer of them, with spaces in between or at least an edge to the metropolitan area. The CBD would be more intensely developed, as would some nodes, but areas farther from subcenters would decline in intensity. Transportation between nodes would be some high capacity form.

---

17 Some empirical evidence indicates that nodes within metropolitan areas arise naturally and conform at least in size if not spatial arrangement to the expected distributions. Pivo (1990) studies subcenters in six cities over time, and Loukissas in TRB (1990) provides descriptive data about Houston, a city with relatively weak planning and zoning. Giuliano and Small (1991) present quantitative information about subcenters in Los Angeles, pointing out that the subcenter sizes approximate the rank-size rule or Pareto distribution. Such a distribution is strongly suggestive of Christaller-Losch kinds of spatial structures. Existing nodal arrangements, however, are spaced and shaped for automobile and truck access, not transit modes.

18 Empirical research indicates that even in a growing area, there are subcenters that decline and disappear into the overall gradient (Giuliano and Small, 1991).
Jobs-Housing Balance. The issue of interspersing employment and residences has figured in transportation and land use discussions since the identification of the density gradient. The first question was the rate at which employment would decentralize as population (residences) spread out, and the answer was fairly quickly. More recently, jobs-housing balance has focused on non-radial commuting patterns and the extent to which land use controls impeded or could encourage shorter commutes through spatially mixed land uses.

In the efficient city, locators are not distinguished with respect to function, such as commercial, residential, and transportation. They are all economic activities choosing their locations and intensity levels according to their land consumption, other inputs, and output market tradeoffs. Assuming externalities are controlled, there is no more or less reason for residence to be located next to residence as next to commercial or transportation. Activities which need to cluster will do so; activities that value land as an input relatively more than access will occupy lower levels of the rent tent.

From this distribution will arise some equilibrium level of average commuting distance. Whether current levels are too (inefficiently) high or too low depends upon whether existing market distortions and imperfections create a bias one way or the other. In fact, the bias is toward longer trips, because transportation pricing (too low), land use restrictions (homogeneous zones), and uncontrolled externalities all result in greater spreading out and less intermixing of housing and employment. The efficient city would be more compact, have more mixed land use, and closer ties between household workplaces and residential locations. This is not to say that workplace location(s) will dominate residence location choice, only that commuting costs will be more important than at present.  

CONCLUSIONS

While each of these effects may be small, by itself, the cumulative impact would be large. Thus efficient cities would look considerably different from current ones, and surprisingly like the ideal cities described by physical planners from another era. Travel would not seem all that much different, but average occupancy levels would be much higher and transit would probably capture a higher mode share. Less national income would be consumed in transportation, and non-users would be relieved of any tax burdens. Urban areas would provide higher environmental quality, and would be more clustered into nodes, but the apparent differences would be less than the real ones. Current overcapacity would be eliminated.

19 Research in urban spatial structure is sometimes diverted into relatively unproductive pursuits by maintaining distinctions between land use categories such as commercial and residential; an example is the "wasteful commuting" debate. For the last word, see Giuliano and Small (1992).
Summary of Changes in Highway System Characteristics

**Delay.** Congestion would still exist, in the sense that average travel times would be greater than free flow travel times, but additional time would be small because of the shape of the average travel time function and the optimum capacity. There would be no “excess” delay, and little congestion of any sort, save for occasional accidents.

**Occupancy.** Vehicle occupancies would be higher. Carpooling, van pools, jitneys, buses, etc., would increase relative to solo auto. Most peak travellers value their time and would remain in the peak, receiving better service, often saving money by carpooling as well. Occupancy would correlate with land use intensity and corridor demand.

**Parking.** Full cost pricing of parking would increase the price to the user, and would have significant impacts on mode choice, congestion, and land use. Fewer vehicles would enter dense areas, and they would go directly to parking rather than cruising for cheap on-street spaces. The supply of parking would be reduced, and its land consumption even more.

**Externalities.** Transportation externalities would be lower. Noise, danger, pollution, and other negative effects could be controlled at less cost than current damages, reducing total vehicle travel somewhat and person travel only slightly.

**Highway Extent.** Highway mileage and land consumption would be less. High demand corridors would shift to more efficient modes or high occupancy vehicles rather than consume more than a third of the total land area, as is common at present.

---

Summary of Changes in Urban Transit

**Fares.** Conventional (fixed route) transit would have fare structures differentiated by time of day of the week, direction, service quality (e.g., seated), and market (both cost and demand factors, to include multi-class service and discriminatory pricing). Overall fares would probably be higher, although some categories could be lower.

**Service.** Route structures would be rationalized, saving substantially more in costs than would be lost in revenues. Overall service quality would be higher (in part from the absence of congestion), resulting in increased ridership.

**Occupancy.** Peak occupancy (crowding) would be less and off-peak about the same or higher (some demand shifted to the off-peak, and less unneeded service would be offered).

**Cost.** Unit capital and operating costs ($/vehicle hour) would be reduced through more efficient production methods and management, and unit output cost ($/passenger trip) reduced through improved deployment and service design.

**Mode Split.** High capacity passenger modes would carry higher shares of total person travel.

**Suburban.** Flexible (non-conventional) transit will expand in suburban areas, once highways and parking are properly priced and nodal intensification has taken place.
### Summary of Land Use Effects

**Nodes.** A metropolitan area would consist of a spatial hierarchy of nodes, with usually one primary node, and numerous subcenters arranged in patterns of interlocking lattices. Nodes would be more intensely developed, smaller in extent (stronger but more limited spatial monopoly), and fewer in number (some existing centers would combine).

**Rent/Intensity Gradient.** The Central Business District and other urban nodes would be more "compact" in the sense of a steeper rent-intensity gradient.

**Size.** The primary CBD will be larger (from eliminating congestion inefficiencies). Dense nodes served by high capacity transportation modes would flatten the metropolitan rent gradient, on the assumption that line haul scale economies would offset reduced subsidy. Efficient cities would be more competitive, economically, hence more desirable business locations.

**Agglomeration.** The reasons people need to engage in movements among particular locations, and are willing to pay for higher access, will not be reduced (e.g., by telecommuting). Interaction benefits (e.g., economies of scope within a node) reduce unit costs and increase the attractiveness of spatial agglomeration.

**Environmental Quality.** Noise, particulate, fumes, smog, venturi effects, and shadow would be lessened, enhancing urban environments. Water quality and groundwater recharge would be improved because compact development permits better control of runoff and pollutants.

**Building Height.** Maximum heights might increase, but greater control of externalities suggests a reduction in average height as well as coverage, resulting in lower buildings with more parks and plazas. This would mean a reduction of the highest density nodes, despite overall intensification of urban land use.

**Transportation Land Use.** Auto vehicle travel would be less, auto person travel somewhat less, walking and bicycle travel higher, parking space greatly reduced, and perhaps more elevator (vertical) travel would occur, resulting in much less land used for transportation.
Summary of Land Use Effects (cont’d)

**Neighborhood Externalities.** Neighborhood effects (light, noise, visual impacts, air and water pollution) would be reduced, allowing closer proximity of mixed uses. Residential neighborhoods would have little through traffic and higher amenities from removal of highway externalities, achieved through a combination of emissions reductions, noise barriers, wider ROW, and traffic control measures.

**Delivery.** Urban shipping, deliveries, trash collection, and similar activities would take place at night (daytime road and parking space would be too expensive).

**Intercity Bus and Rail.** Bus service would be more differentiated (no frills luxury, Las Vegas express), but low cost (price). Passenger rail service will link major urban areas (over 500,000) for distances of 50-300 miles. Terminals would be in urban centers, accessible to transit.

**Growth Control.** The marginal resident or employer will face the full marginal cost of its location decision, including urban infrastructure and environmental costs. Infrastructure will be largely paid by users as depreciation occurs. Hence marginal costs of new residents would not increase sharply; to the extent they do, all residents and employers (not just new ones) should face the same marginal cost and decide to leave or stay. The effect will be to keep cities from becoming excessively large.

**Transportation Share of GDP.** Transportation’s share of resources would diminish, both capital infrastructure and operating costs. The efficient system would be smaller than the existing one.

**General Taxes.** To the extent that non-distorting taxes (such as property and sales taxes) to support general government are required, they would be imposed neutrally on transportation as well as other land uses and economic activities.

Equity Impacts

**High Occupancy Modes.** Improvement of high occupancy modes creates benefits for low income travellers (e.g., bus users) at the expense of high (e.g., peak auto).

**Peak versus Off-Peak.** Marginal cost pricing will benefit off-peak (time and direction) travellers, who generally have lower than average income. Peak travellers (auto commuters, commuter rail) will see the greatest increase in price.

**Indirect Taxes.** Full cost pricing will reduce direct and indirect taxpayer subsidies, which in the aggregate are more regressive than the distribution of efficient user charges. In addition, transportation activities would generate surpluses which could be tapped for revenues to support general government (i.e., property taxes levied on highways).

**Housing.** Inefficiencies in the housing sector (interest and tax deductions, credit subsidies, homogeneous zoning) complement transportation inefficiencies to favor middle and upper income housing. Removal of this bias, combined with efficient control of neighborhood effects, would permit moderate and high density mixed-use and mixed-income housing construction.
REFERENCES


Register, Richard, “What is an Ecocity?” Earth Island Journal, Fall 1987, pp. 31-33.


