nferences
s.

LOCAL SERVICE PRICING POLICIES AND THEIR EFFECT ON URBAN SPATIAL STRUCTURE

Edited by
Paul B. Downing

Published for
THE BRITISH COLUMBIA INSTITUTE FOR ECONOMIC POLICY ANALYSIS

UNIVERSITY OF BRITISH COLUMBIA PRESS
VANCOUVER

This book is the second in a series based on the Economic Policy Conferences of the British Columbia Institute for Economic Policy Analysis.

Conference on the Pricing of Local Services and Effects on Urban Spatial Structure
Vancouver, 1974

When To Build What

MASON GAFFNEY

INTRODUCTION

This paper purports to solve a particular kind of problem that characterizes urban expansion and evolution: when to replace a collection of individual apparatuses (CIA) with a mass system. Examples include replacing individual septic tanks by sewers, wells by public water supply, private cars by mass transit, trash burners by public pickup, coal or oil fuel by line-distributed gas or electric power, individual by community antennae, individual driveways by alleys or a subdivision, individual roadside business signs by collective locational advertising, individual messenger by postal service, individual deliveries by united parcel service, tank trucks by an oil pipeline, individual stores by a shopping centre, basement pumps and periodic cleanups by storm sewers and flood control, and so on. On a larger scale one meets the same problems in replacing a collection of small, Balkanized local sewer systems by a metropolitan or regional system.

A feature of most mass systems is a distributive network using street space and other rights of way to link individual sites to a common load centre. (The streets themselves are also a sort of mass distributive network, although not self-contained.) Hence I offer the term "intersite" to distinguish these distributive lines with their peculiar cost traits and relation to land from "onsite" works. We may also distinguish a street from an offstreet sector, though I will use the terms interchangeably.

Mass systems often include an offstreet load centre. But distribution costs (onstreet, offsite) generally exceed load centre costs, especially in this age of urban sprawl with inflated costs of intersite linkage. And the street costs give the mass system their distinctive cost characteristics in relation to volume of service and area serviced.

The outstanding cost trait is long-run decreasing costs as volume of service rises. This trait is an empirical fact of such wide currency that here I merely postulate it.

Many analysts overlook, however, the essential source of decreasing costs. The decreasing costs are realized only by increasing volume inside a given area, not by expanding the service area and lengthening the lines. On the contrary, lengthening lines means carrying some units of service farther from load centre to user. Mileage rises, and average cost per mile rises, too, if density falls, as it well may. A counterpart of decreasing cost to volume is increasing cost to distance.

This contrast between volume effects and distance effects is central to our subject, meshing it closely with urban land economics. An example showing the practical force of the point has been worked out by Paul Downing.[1] This work is so telling and pointed that I reproduce here in full his table (Table 1).

While offstreet load centres may also be in a stage of decreasing costs to additional scale, they may not. Obviously they enjoy some such economies, which is why it pays to link them to sites via an expensive intersite network. But, typically, scale economies in a load centre are fully or largely realized, while intersite economies are not. Thus a major metropolis has multiple water plants, sewage treatment plants, power plants, trash disposal sites, and the like, each plant often consisting of more than one parallel, duplicate bank, line, or other facility. Water supply acquisition often requires longer aqueducts, larger reservoirs that fill less often, use of previously submarginal small streams, and other operations of increasing cost. Larger landfills require land farther out, multiplying the hauling cost. The scale economies to be realized by further growth are primarily in distribution or collection, that is, our "intersite" sector. They are economies of density, not of over-all system volume.

This characteristic of the intersite sector, and its contrast to the onsite sector, was brought out by P.A. Stone's study of housing costs and density in England.[2] Stone distinguished street costs from offstreet costs and found the former fell (per unit, of course) with higher density, while the latter rose. Putting this another way, a builder crowding more dwelling units on the same land meets diminishing returns on the site, while still enjoying increasing returns to the capital needed to enlarge street improvements.

In contrast to mass systems, CIA's do not enjoy economies of density. The individual apparatuses usually are too small, their size being limited by the demands of one household. But it would be a fallacy of composition to conclude the collection is too small. Further scale economies of the whole are offset by negative externalities. Thus, larger wells are more efficient to the individual, but in concert they may lower a water table; larger septic tanks saturate a seepage bed; larger autos worsen street congestion and air pollution; larger competing billboards and come-on signs neutralize each other and degrade a neighbourhood; and so on. So a CIA at high density, viewed as a whole, often suffers from increasing social costs, even though its individual components each seem to be in a stage of decreasing costs.

TABLE 1

MARGINAL COST OF SEWAGE COLLECTION AND TREATMENT WITH DISTANCE AND DENSITY-LOW ESTIMATE
(1957 to 59 dollars)

Density (People/Acre)	Annual Cost Category ($/Capita/Yr.)	Distance from Subdivision to Treatment Plant (Miles) 5	10	15	20	25	30
0.4	Collection	33.60	33.60	33.60	33.60	33.60	33.60
	Transmission	4.50	16.20	28.80	41.30	53.80	66.20
	Treatment	2.07	2.07	2.07	2.07	2.07	2.07
	Total	40.17	51.87	64.47	76.97	88.47	101.87
1	Collection	14.59	14.59	14.59	14.59	14.59	14.59
	Transmission	1.80	6.50	11.50	16.50	21.50	26.50
	Treatment	2.07	2.07	2.07	2.07	2.07	2.07
	Total	18.46	23.16	28.16	33.16	38.16	43.16
4	Collection	6.46	6.46	6.46	6.46	6.46	6.46
	Transmission	0.80	30.80	6.80	9.80	12.80	15.80
	Treatment	2.07	2.07	2.07	2.07	2.07	2.07
	Total	9.33	12.33	15.33	18.33	21.33	24.33
16	Collection	4.86	4.86	4.86	4.86	4.86	4.86
	Transmission	0.50	3.10	5.70	8.30	10.90	13.50
	Treatment	1.07	2.07	2.07	2.07	2.07	2.07
	Total	7.43	9.93	12.63	15.23	17.83	20.43
64	Collection	1.22	1.22	1.22	1.22	1.22	1.22
	Transmission	0.70	2.30	4.60	7.10	9.60	12.10
	Treatment	2.07	2.07	2.07	2.07	2.07	2.07
	Total	3.99	5.59	7.89	10.39	12.89	15.39
128	Collection	0.62	0.62	0.62	0.62	0.62	0.62
	Transmission	0.60	2.10	3.90	5.50	7.10	9.70
	Treatment	2.07	2.07	2.07	2.07	2.07	2.07
	Total	3.29	4.79	6.59	8.19	9.79	12.39
256	Collection	0.27	0.27	0.27	0.27	0.27	0.27
	Transmission	0.60	1.75	2.90	4.05	5.20	6.35
	Treatment	2.07	2.07	2.07	2.07	2.07	2.07
	Total	2.94	4.09	5.24	6.39	7.44	8.69
512	Collection	0.16	0.16	0.16	0.16	0.16	0.16
	Transmission	0.55	1.65	2.75	3.85	4.95	6.05
	Treatment	2.07	2.07	2.07	2.07	2.07	2.07
	Total	2.78	3.88	4.98	6.08	7.18	8.28

Also, since our subject is the replacement of existing apparatuses of given capacity, these may suffer short-run rising costs regardless of the long-run possibilities.

Thus, when density approaches a level at which we consider replacing a CIA with a mass system, it is a question of replacing an increasing cost system with a decreasing cost system. This poses some difficult social and political questions, the answers to which determine much about the character of cities and society.

DISTINCTIVE FEATURES OF THIS REPLACEMENT DECISION

The basic criterion of all replacement timing is to replace when the rent of the replacement, or challenger (net of capital costs), exceeds the quasi-rent of the sunk facility, or defender (gross of historical capital costs). This basic criterion holds here, but with distinctive wrinkles.

When an older offstreet building is replaced by a new one, there may be a jump in total floor space, occasionally a large jump. But the increase is limited by increasing costs, since more capital on one site meets diminishing returns. It is normally this cost behaviour, rather than a demand constraint, that limits the size and rents of an offstreet challenger.

When the challenger is an intersite facility enjoying decreasing costs, the limit on its scale and rents must be on the demand side. If demand is elastic the increase of volume may be very large. Demand affects rents not only by affecting total volume, but also by affecting the rents generated by intramarginal volume. Much or all of the rents of intersite works are consumer surplus, which in turn is captured in offstreet rents. The intersite works themselves may show a specious deficit, when valued on the basis of a fare-box or user charge. The full rent of a mass system challenger, as properly used in a showdown, should include the challenger's contribution to offstreet rents, which may be its only positive rents, yet are often high enough to justify its use.

So the demand schedule assumes a key role in the decision. But demand is largely (not totally) a function of offstreet density; and density in turn is controlled by public policies such as zoning, assessment procedures, and income tax loopholes. Ultimately it is limited by increasing costs to offstreet density. That is, diminishing returns of capital applied to building sites are still a limit on the intersite sector, but reveal themselves in the form of diminishing demand rather than of rising costs. These then are the elements of our problem.

If demand were inelastic, there could be no big rise of volume. In this case, the way to maximize rent is simply to minimize cost at the fixed volume given by inelastic demand. Replacement then is timely when the vertical demand schedule, drifting to the right for reasons other than price, reaches the volume where the rising average cost of the CIA intersects the falling average cost curve of the mass system.

But demand is elastic, for two reasons. The lesser one is that usually given sole weight in studies of demand for utility service, to wit, increasing use per capita. Even on this limited basis, researchers have found great price sensitivity among users of water, power, mass transit, and so on.[3]

The larger reason, usually overlooked, is immigration resulting in higher density. Remember we are discussing a mass system serving a given area which will naturally attract new high-density development when lots with cheap services are made available. When we let density vary, demand becomes extremely elastic. In analyzing urban growth and successive land uses, this is a realistic assumption; the transition to higher density is the central theme. Thus this replacement decision is epochal for a neighbourhood, changing its whole density and character, entailing replacement of much offstreet capital which becomes locationally obsolete as mass systems replace the CIA's. In addition, the higher density appropriate to one mass system tips the balance for others, a form of interdependency so close that ideally the entire conversion decision should be treated as a unit. Although this last point is too ambitious for this article to treat, the analytical framework presented here could be adapted and elaborated to that end.

So the distinctive aspect of our replacement problem is the large difference in optimal volume between the two alternatives, a defender with increasing costs and a challenger with decreasing costs. This is a discontinuity in economic development. It is not a decision that can be approached tentatively or incrementally on little cat feet. No replacement decision is, but this one is more shaking than most. It means a leap into greater volume; it is social; it changes the economic signals for all site owners and, through their responses, for all other intersite distributors.

Just as the epochal nature of this replacement might be masked by assuming inelastic demand, so might it be damped (although not eliminated) by a policy of average cost pricing. However, marginal cost pricing is the ideal of economic theorists as well as a goal sought—and often approached—by practical administrators. Indeed, it is the only kind of pricing consistent with the replacement criterion of selecting the system of greater rents, since average cost pricing fails to maximize rents anyway. So I will begin by defining criteria of replacement timing under ideal marginal cost pricing with price equal to marginal social cost. This accentuates the jump in volume.

Note this implies that a user charge is imposed on the CIA equal to marginal external damages, while the mass system is partially financed from flat rate charges (on beneficiaries, preferably) to the end that user charges fall below average costs.

A MODEL OF OPTIMAL REPLACEMENT

The CIA defender, with rising average social costs, suffers marginal social costs above its average. The mass system challenger enjoys long-run marginal costs below its average. Figure 1 shows the relationship graphically.

Note in Figure 1 how the demand curve crosses a marginal cost curve twice: the short-run marginal costs of the defender (SMC_D) at volume A, and the long-run marginal costs of the challenger (LMC_C) at the greater volume *B*. This shows a great difference in the optimal volume under the two systems, each crossing being optimal for its respective system.

Which is better? The decision goes to that yielding the higher net rent where rent is total benefit less total cost. One way to find this rent is simply to sum up benefits and costs at *A* and *B*. A second way is to calculate the increments to cost and benefit between volumes *A* and *B*. Method two has an advantage in that it is purely comparative, requiring no estimate of total consumer surplus. I will therefore develop method two.

Let us label the gross increment to cost between *A* and *B* the discrete marginal cost (*DMC*), and the gain of welfare correspondingly as discrete marginal benefit (*DMB*). Replacement becomes timely when *DMB* > *DMC*.

DMB is the area under the curve of demand, or marginal benefit, from *A* to *B*.

DMC is the excess of LTC_C at *B* over SVC_D at *A*. Visualize this as the difference of two rectangles based on the origin as the southwest corner. The smaller rectangle is determined by $SAVC_D$ at *A* as its northeast corner, the larger rectangle by LAC_C at *B*.[4]

Figure 2 is a graphic method of showing the differences between *DMB* and *DMC* by the familiar device of cancelling overlapping areas common to the two and shading in areas not common. Slanted shading shows clear benefits; horizontal, clear costs.

In Figure 2 there is a net gain showing replacement is overdue. (A similar layout on Figure 1 would show the reverse.) Note, incidentally, that in neither case would the challenger show a positive fare-box or user charge rent under average cost pricing, or any uniform pricing, the marginal demand always lying below average cost. The reasoning here is based on counting consumer surplus with benefits—a process whose plausibility is more evident if we say it compares average cost with average benefit.[5]

Inspection of Figure 2 suggests a rough rule of thumb. If the demand curve is straight and $SAVC_D$ at *A* is fairly close to LAC_C at *B*, then added

FIGURE 1
COSTS (AVERAGE AND MARGINAL) OF DEFENDER AND CHALLENGER, WITH ELASTIC DEMAND SCHEDULE

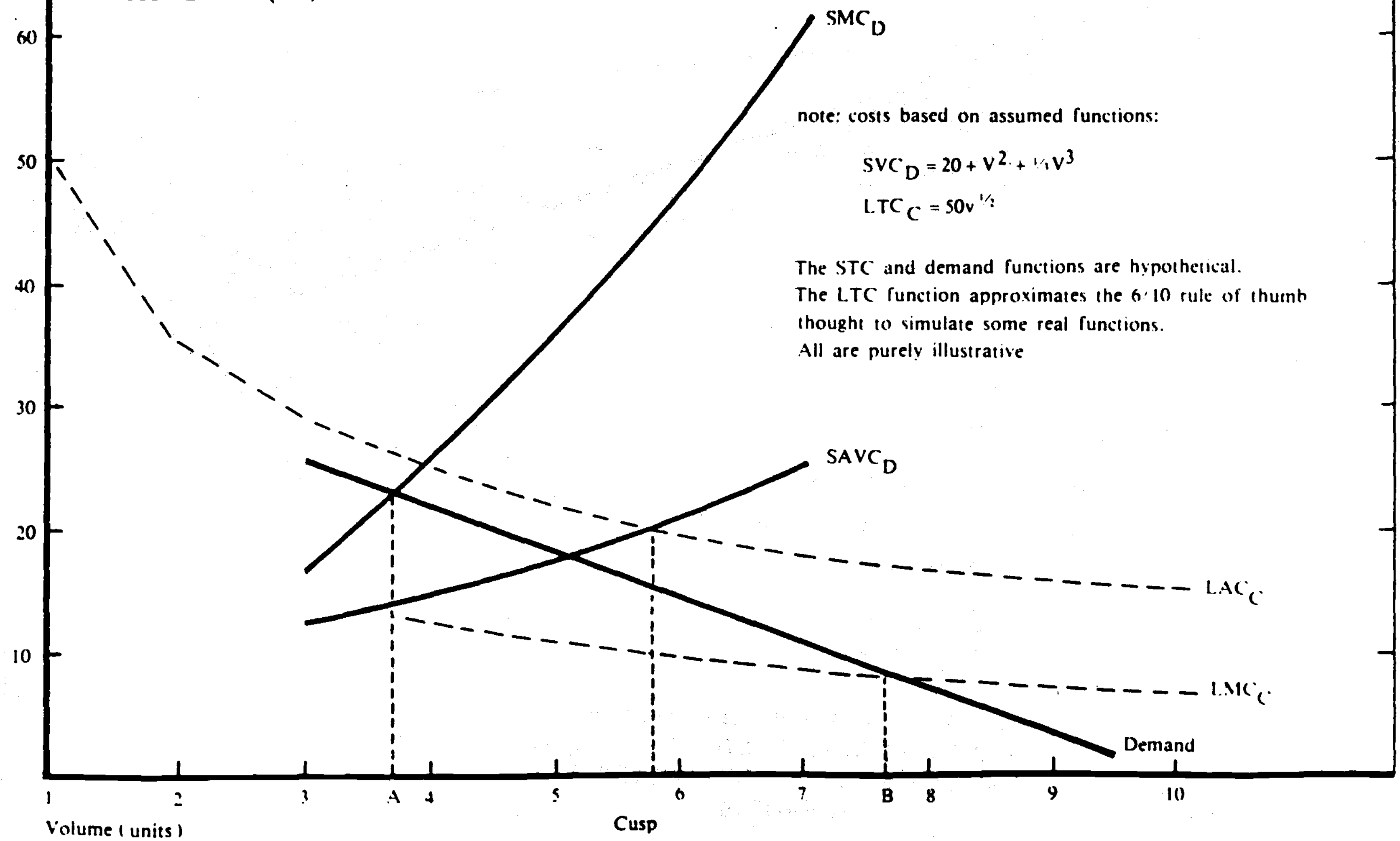
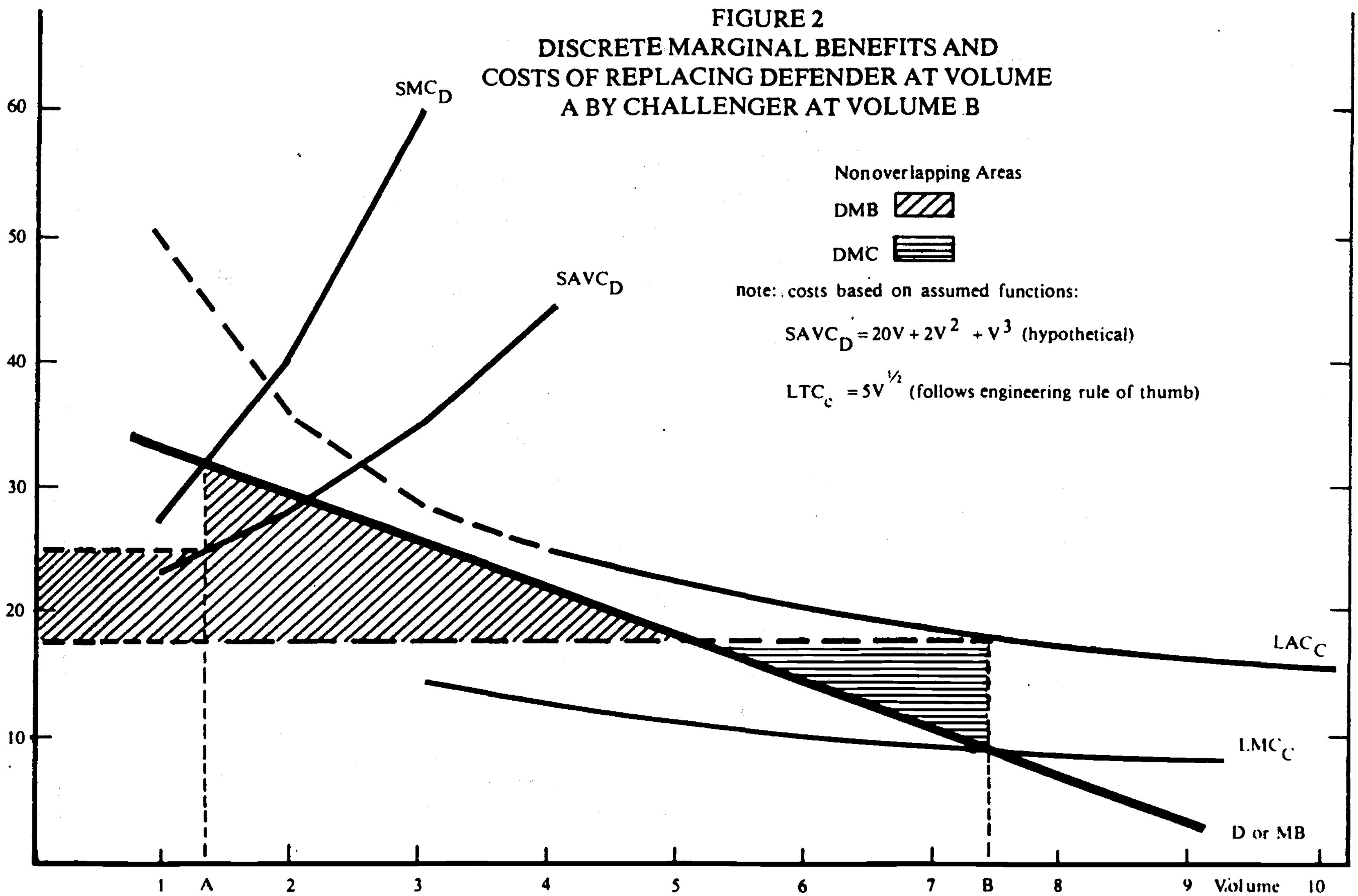


FIGURE 2
DISCRETE MARGINAL BENEFITS AND
COSTS OF REPLACING DEFENDER AT VOLUME
A BY CHALLENGER AT VOLUME B



benefits and costs registered on the graph from A to B are about a stand-off, so the balance hinges on the DMB shown left of A. This in turn depends on whether any decline of AC is achieved. The rule of thumb is if average costs can be lowered and the necessary added volume sold at marginal cost, it is time for a closer look to see if it is time to replace. Obviously this rule may be wrong if the marginal cost-average cost relations are much different from those shown.

A useful concept now is the ratio of DMC to the added volume, B-A. Call this discrete marginal unit cost ($DMUC$). It is the mean cost of serving the added volume. Anytime AC_C at B is less than $SAVC_D$ at A, it means the added volume is supplied more cheaply by introducing the mass system in the given area being considered than it would be by maintaining the CIA in a competing location. *A fortiori*, the mass system adds the volume more cheaply than would duplicating a CIA in a virgin area and thus incurring capital or fixed costs as well as variable ones.

This gives a clue to the strength of the forces making for an elastic demand through immigration. The level of demand is not going to fall much below the full cost of servicing alternative sites. An increment of population can often be serviced more cheaply by raising density in mass-serviced areas. Of course a resistance to density may limit this truth, but there are also gains to aggregate neighbourhood density, other than those treated here, which focus demand on an area precisely because it is already thickly settled. That is, diminishing returns to buildings on individual sites cannot necessarily be translated into diminishing returns to neighbourhood density, if there are strong positive external linkages, such as characterize social and economic life.

UNDERSCORING SOME MORALS

Using the analytical framework sketched out, let us now explicate some important lessons and inferences.

a. The advantage of a mass system is to be found in larger volume, not simply in lower costs at fixed volume. This is more than "replacement"; it is expansion. No analysis couched purely in minimizing costs is adequate. To optimize is to maximize rents.

b. Marginal cost pricing favours mass systems over CIA's. Average cost pricing would defer replacement until the demand curve touched the LAC_C below $SAVC_D$. If the demand curve were steeper than the LAC_C curve, that would not occur until the cusp where the two AC curves intersect.

   Conversely, failure to use marginal cost pricing favours CIA's. If one accepts marginal cost pricing as the rational economic ideal, failure to use it constitutes an uneconomic bias against mass systems.

c. Marginal cost pricing lengthens the jump in volume between systems and accounts for most of the discontinuity. The difference in marginal costs is much greater than that in average costs.

   Note that the discontinuity between optimal volumes of defender and challenger is not due to any discontinuity of the cost functions. The curves as drawn presuppose a complete range of alternative sizes, and in practice this is approximated; pipes and wires come in a full spectrum of diameters and gauges. So there is no technological straight jacket to account for the discontinuity. The explanation is in the interplay of cost and demand functions, as shown.

d. High elasticity of demand is essential to a jump between systems. Elastic demand is assured by immigration, raising density.

   There is a problem of lag of demand behind building the mass systems. Perfect synchronization is impossible whenever the mass system is prerequisite. Clearly the responsibility to lead is the mass system's. This lag adds to costs, however, a matter developed below.

e. The textbook method of price discrimination with quantity discounts, as practised by utilities universally, will not work as an approach to marginal cost pricing in this case. Immigrants will not necessarily buy much more per meter. The gain is in larger volume per acre, rather than per meter. A flat rate per acre or per dollar of land value is the more economical way to meet the deficit of marginal cost pricing. This requires imposing a tax.[6]

f. The city should not taper gradually into the country. Ideally there should be quite a sharp edge between a serviced area with mass systems and a CIA area at lower density. This must be modified because there are several mass systems at different cost levels. Power grids outreach sewers, for example. On the other hand, a universal reliance on high density makes the different systems complementary and reinforcing in any given area.

g. Optimal public policy requires recognizing the negative externalities of CIA's and the positive externalities of mass sytems. Our rising CIA cost curves reflect these externalities—the individually incurred costs are often falling. In regard to mass systems, the cost curves reflect no externalities, but the feasibility of marginal cost pricing hangs on recognizing consumer surplus as an externality picked up in rents of the land serviced below average cost.

h. Mass systems which attract and serve immigrants are not prodigal of capital even though they may require a large lump of it. By attracting immigrants they obviate serving them elsewhere at higher cost.

INSTITUTIONAL BARRIERS TO MASS SYSTEMS

There are several institutional barriers to mass systems that prevent society's realizing to the full their inherent savings. Most of these reflect the age-old bias of those people who live at lower density towards those who, from personal choice or budgetary constraint, live at higher density and degree of interdependency. It is the bias of the Indian against the European, of the hunter against the stockman, of the stockman against the sodbuster and pig farmer, of the grain grower against the irrigator, of the farmer against the subdivider, of the suburbanite against the city-dweller, of the detached householder against the apartment-dweller, of the resident against commerce. We might even say, it is the bias of Cain against Abel. In recent times, the wealth and overrepresentation of suburbanites and farmers have given the low-density forces power to hobble mass sytems. We are all familiar with the issue and the powerful emotions it arouses. I take no position on whether one density has more social merit than another. The disqualifying weakness of most assertions on optimal density is that the asserter seeks to impose his own tastes on, or impute them to, others. Often he is dissembling his misanthropy, ethnic prejudice, class interest, callousness to poverty, and taste for capital gains under a cloak of naturalism.

Another source of bias is the indecisiveness of people faced with a complex social transition. They ease into it incrementally, feeling their way, tolerating and even welcoming checks and balances. The problem here is that the case happens to be one where optimal transition is not gradual but abrupt, not groping but bold.

a. Among the worst biases against mass systems is that the specific people who live at high density in mass systems do not receive lower prices or rates to reflect the social savings the higher density affords. Utility rates seldom reflect this factor. Residential densities in a typical metropolis vary from fifty dwelling units per acre down to one-fifth, that is, one two-hundred-fiftieth of the maximum, and all pay the same rates. In Milwaukee, the most densely peopled 3 per cent of the residential area houses some 25 per cent of the people; the most expansive 10 per cent of the people occupy over 50 per cent of the area, further from load centres. Yet all pay the same utility rates.

Referring back to Figures 1 and 2, we can see that the gains shown there are posited on cost savings from higher demand inside fixed areas being passed on to buyers in lower prices. Herein lies much of the incentive to increase density. But in practice the cost savings are diluted through a large system and enjoyed equally by low-density dwellers. So the incentive to achieve the savings is destroyed.

Price discrimination as practised today cannot help. Quantity discounts based on the meter do not reward higher density; neither do they reward early improvement of land, which minimizes the sterile waiting period before a customer buys anything at all.

Often, indeed, discounts based on meter volume react perversely and penalize higher density. Suppose a block contains one spreading old business at low density but with enough volume to qualify for low rates. The owner subdivides, and twenty smaller firms move in. Volume per block rises, but volume per meter falls, so average rates rise.

This kind of price discrimination may make sense to a monopoly firm exploiting its power. It does not achieve the social goals of marginal cost pricing, however, unless, in fact, the costs of metering and billing are paramount. One hears such premises, but they are not factual. If they were, the meter industry would be a giant. Table 2 is a cost breakdown assembled from questionnaires answered by 188 water supply facilities and 157 wastewater disposal facilities in the United States for 1966. "Meters" are only 2.3 per cent of the water supply costs and are not itemized at all among disposal costs, presumably because water meters do double duty. The lion's share of water costs are transmission (31 per cent) and mains (27 per cent); of disposal costs, interceptors (39 per cent) and trunk sewers (25 per cent). I believe most economists do not have any notion of the degree to which urban sprawl has inflated line costs relative to other costs, even treatment plant or other load centre costs. It almost seems mean to add that the cost of reading meters varies with the distance between them, as does the hidden social cost of delivering bills by United States mail. It is less mean to note that a given number of customers must be metered wherever they are, so the higher metering costs in the subdivided block are offset by lower costs elsewhere. The net change is a shortening of lines.

Actually, with meter costs so small relative to line costs, a utility gains by having volume of any given amount split up among many customers. This is because the pooled demand of twenty individuals is much steadier than the demand of one individual. Load factors rise, peaks are levelled, and unit costs fall.

b. Equally biased are public policies aimed against intensive land use. Some would-be dictators of public taste excoriate density as a plague and paint subdividers and developers as conscienceless fast buck artists raking in the green by exploiting the tolerance of their cultivated and restrained neighbours. Many policies reflect this view:

1. Low-density zoning and height limitation are the most obvious. Legal protection of low-density covenants is similar.

2. The property tax on buildings is equally biased. By adding to building costs, this tax motivates substituting land for capital; that is, horizontal spread.

3. Administration of the property tax on land values is generally, albeit extralegally, biased against density. Intensification of use is made the occasion to reassess land whose market value and use potential have appreciated over the last twenty-five years. Thus an owner who subdivides a farm and lays water and sewer lines finds his assessment higher, not merely reflecting his recent outlays but also the unrealized increments of

TABLE 2

PER CENT DISTRIBUTION OF COST AMONG SYSTEM COMPONENTS, 1966

Water Supply	%	%
Acquisition	38.7	
Source of supply		7.6
Transmission		31.1
Distribution	35.5	
Mains		27.1
Booster stations		3.0
Storage in distribution system		3.9
Hydrants		1.5
Treatment	14.3	
Screening		0.5
Chemical handling		2.6
Coag. & softening		2.1
Sedimentation		2.2
Filtration		4.6
Disinfection		0.8
Plant storage		1.5
Meters	2.3	
Unassigned	9.1	
Pumping		3.9
Lab. equipment		0.1
Flow measurement		0.1
Instruments & controls		0.6
Other		4.6
Wastewater Disposal		
Collection	69.1	
Interceptors		39.1
Trunk sewers		24.7
Force mains		3.6
Lift stations		1.7
Treatment	28.5	
Pumping		2.1
Motor		0.4
Valves & lines		2.3
Screening		1.1
Grit removal		2.1
Sedimentation		5.4
Trickling filtration		0.4
Activated sludge		6.2
Digestion		3.9
Sludge dewatering		2.0
Sludge handling & disposal		1.9
Disinfection		0.1
Unassigned	2.5	
Lab. equipment		0.1
Flow measurement		0.1
Instruments & controls		1.1
Other		1.0

Source: *Water and Wastes Engineering*, April, 1966, pp. 34-39.

the generation past. On the other hand, if he installs a new well or septic tank there is no occasion to reassess because there is no change in use class. Also, since the market value of the land is based on future urbanization, these CIA improvements are obsolete immediately and add nothing to value.

4. Income tax loopholes for increments to land values subsidize and encourage holdouts.

c. There are hardly any user charges levied to make individuals feel the external social costs of CIA's. A rarity is Orange County, California, which charges water well pumpers for lowering the water table. Generally one may pump without charge, pollute aquifers and lakes and streams, issue loud noises from gas motors, pollute the air, congest streets with cars and airports with private planes, degrade highways with billboards, and commit nuisance after nuisance without charge.

So the costs perceived by individuals are well below the social costs. This is particularly true for marginal costs, inasmuch as the private, internalized component of the total cost is decreasing. It is the external nuisance that makes for increasing costs.

It does not follow in every instance that the mass system is deferred. External costs are noticed by their victims, and they may organize to act against them. Thus California irrigation districts in areas of falling ground water distribute surface water below cost with the purpose of discouraging pumping. This is the exception, however, partly explicable by the fact that appropriative law is likely to encourage wasting water. Our polluted, degraded environment testifies that society often accepts gross damages from CIA's before organizing effective resistance or alternatives.

The model of optimal replacement developed in the third section of this article assumes that optimal user charges are levied on the CIA. Where there are no such charges the increased costs (*DMC*) of the mass system will be less than shown since the CIA will have grown far into the area where its marginal social costs exceed marginal benefits.

On the other hand, there is a danger of overstating increased benefits (*DMB*) by inadvertently assuming that existential marginal social cost equals demand price for the present volume and so gives us a point on the demand curve. It is the much lower marginal individual cost that reveals what people will pay for more volume. There are many cases where applying user charges to optimize the behaviour of CIA's would be more economical than going to a mass system. We should be wary of the California irrigation districts, for example, which waste surface water to save ground water. We should look sharply at a mass transit analysis that states benefits in terms of reduced costs of auto congestion.

These costs may be above the demand price anyway; we might be happier simply driving less than travelling more by subway.

However, when we consider that immigration is the prime source of demand elasticity in any given area, the danger of overstating *DMB* recedes. The external social cost of leaky septic tanks, for example, takes the form of discouraging immigration to an area, which in turn reduces the demand for sewer service. Supplying sewers would create its own demand. The same is true of trash pickup to combat littering, or public health measures to combat disease and to obviate the cost of individual doctors.

In addition I harbour an inchoate feeling that the willingness of B to absorb the external damages of A represents a price B pays to do as he pleases and pay no user charges as he damages A. It may be then that the demand price is higher than individual marginal cost and includes some of the spillover damages. The notion is worth exploring, although I do not guarantee results. It is related to the currently popular topic of "option demand." That is, there is an insurance value for me in having water or bus service available that I seldom or never use.

There is, besides, a value for me in having my cleaning woman ride the bus which I never board; certainly the merchant gains from his customers' travels. With these points gathered together, a simple demand curve understates social benefits from added capacity. This will vary with the service in question, of course, and must be evaluated case by case. But increased land rents are a truer guide to social benefits of some mass systems than any demand curve, and they may not be any harder to measure.

Perversely, when antipollution measures are introduced, they generally strike the mass system first. A city sewage plant is large enough to identify, and to sue, and to police, while one-hundred-thousand septic tanks are something else again. Large apartment incinerators and power plants can be noticed and served orders; but four million Los Angeles autos, the major polluters, are a moving target, elusive as Rickover's roving sub, and far more numerous. Antinoise compaigns are focusing on jet aircraft and will doubtless score there first while the ubiquitous flatulent motorcycle and power mower blat on unchallenged.

d. There is too little recognition of the positive externalities of mass systems. It is nearly universal to regard meeting the deficit of mass transit, for example, as a "subsidy." Maury Seldin[7] has a reverse concept of subsidy which I believe more truly gets to the realities. Failure to collect some form of payment from owners of land that has been provided with mass facilities by the local government is "in effect, public subsidy of land speculation. . . . To the extent that the taxes on the land alone are not providing a 'fair share' of the cost in fixed investment in excess capacity of community facilities, the owner of the land is being subsidized. . . ." Amen.

But time and again we hear and read that mass systems must earn a profit and not ride on the backs of suffering taxpayers. The gold star goes to "profitable" water or municipal power systems that "relieve" local taxpayers. One cannot slog through the popular or professional literature on this subject without becoming convinced that most thinking, and therefore presumably most decisions, are based on models of small, private, offstreet firms quite unlike the intersite sector about which decisions are being made.

Once again, there are exceptions. Since loopholes in the federal income tax laws virtually exempt land value increments from taxation, wealthy people prefer to receive income in this particular form and are very sensitive to factors than enhance land values. Furthermore, local property taxes to finance municipal mass systems are deductible from personal income taxes. These are powerful forces to overcome the traditional antipathy of large landowners to mass systems.

A problem is that these motives operate unevenly, varying with the marginal tax brackets of the landowners. These, in turn, vary inversely with density, so the boost in motive is greatest where least appropriate, in low-density suburbs and exurbs. The practical outcome is often that the disproportionate political influence of wealthy speculators is exerted to secure extension of urban mass systems, heavily subsidized by general taxpayers and city utility users. Higher user charges in central areas help drive people outward, justifying yet further extension. The result is premature growth of mass systems in remote suburbs, but not on any systematic economic basis.

e. Mass systems, which should be financed from taxes, are instead favourite objects of taxation. They are not only taxed, but overassessed for property tax—every study of discrimination shows utilities leading the list of overassessed property. This is passed on in higher rates. Excise taxes are loaded on this base.

Many CIA's, on the other hand, are undertaxed. Automobiles escape most property tax assessments, and their accessory land uses are heavily favoured because they need few buildings and land is underassessed. Parking is subsidized everywhere, and cars allowed to invade more and more places.

Billboards enjoy the same tax benefits as parking lots. Both water wells and billboards are undertaxed, for different reasons. I could go on.

f. Mass systems are attractive targets of unionization. Since transit and utilities are monopolies whose services are nonstorable, the social costs of strikes are extremely high. The paralysis of bureaucracy and low politics are visited on mass systems more than on CIA's. Society must not let mass systems be milked or throttled if it is to enjoy their technical benefits.

g. Introduction of different mass systems is imperfectly synchronized. To some extent, land subdivision is a key event that synchronizes layout of distribution networks in given areas, but everyone with experience in this field can recount horrors of bad coordination. It is bad enough when the sewer men rip up the street the week after power men have closed it. Worse, in general, is the effect on demand projections dependent on offstreet density. When mass system capital outlays are synchronized, each can confidently be premised on higher density and demand, and earlier development of demand, than if one goes in alone.

h. In lieu of user charges on CIA's many suburbs adopt low-density zoning which then rigidifies and prevents immigration when demand is ripe for mass systems.

More than inertia can be blamed here. The central problem is schools. Because of low-density zoning, schools cannot react flexibly to an influx of immigrant children. School resources are not invariant. Public schools are a mass system, and they enjoy economies of scale, according to James B. Conant's studies. School busing, chauffeuring, and commuting certainly involve economies of density. But local zoning attitudes are dominated by two other factors.

One is short-run congestion of fixed plant. children stay in any school only three to four years, and their parents' horizon of concern is not much longer.

More fundamental is the redistributive aspect of school finance. Taxes vary with property; social costs vary with family size; benefits vary with intellectual bent and professional interest. So zoning is designed to repel families whose property value per child falls below a certain average. Whoever said that North America idolized motherhood never looked into suburban zoning. Politically, those with large property tend to join with those of low intellectual bent to oppose schools. This is a hard combination to beat.

That all follows from states and provinces requiring localities to make schools common at local expense and then delegating zoning power to localities. The tragedy is that the combination of institutions makes schools, a mass system with decreasing costs, appear, to the localities, to be a fixed resource that needs to be protected from congestion. Their zoning defence against immigration of children not only is bad for children and for schools but for all mass systems that require high density.

SOME PROBLEMS OF CAPITAL COSTS

a. Uneven age of CIA individuals. So far we have treated the CIA as though all members were old and largely depreciated. In fact they are likely to be of uneven age because few areas are fully developed all at once. Some will need early replacement when others are brand new. At best they fail individually for their separate reasons. This is one of the costs of spotty urban expansion.

Part of the current cost of maintaining the CIA therefore is the capital cost of replacing its failing members. This might seem to weight the balance unfairly, but anyone who has nursed an old car understands the problem: someday it has to be junked in spite of new radiator and gas pump.

This factor can add materially to CIA costs. For a basis of estimate let us assume the uneven ages of CIA members are perfectly staggered and all have the same life (L). Every year, $1/L$ individuals are replaced. The yearly cost is C_0/L, where C_0 is the capital cost of the entire CIA.

This is nearly half of what we would have to assign for level amortization of C_0 over L years if we were building the entire CIA new with a full life ahead. The level amortization factor is the capital recovery factor (CRF),

$$\frac{i}{1-(1+i)^{-L}}$$

(see Appendix 1, Table 3, and Figure 3). If L is 25, so $1/L$ is 0.04, the factor is 0.09 (if i is 0.08.)

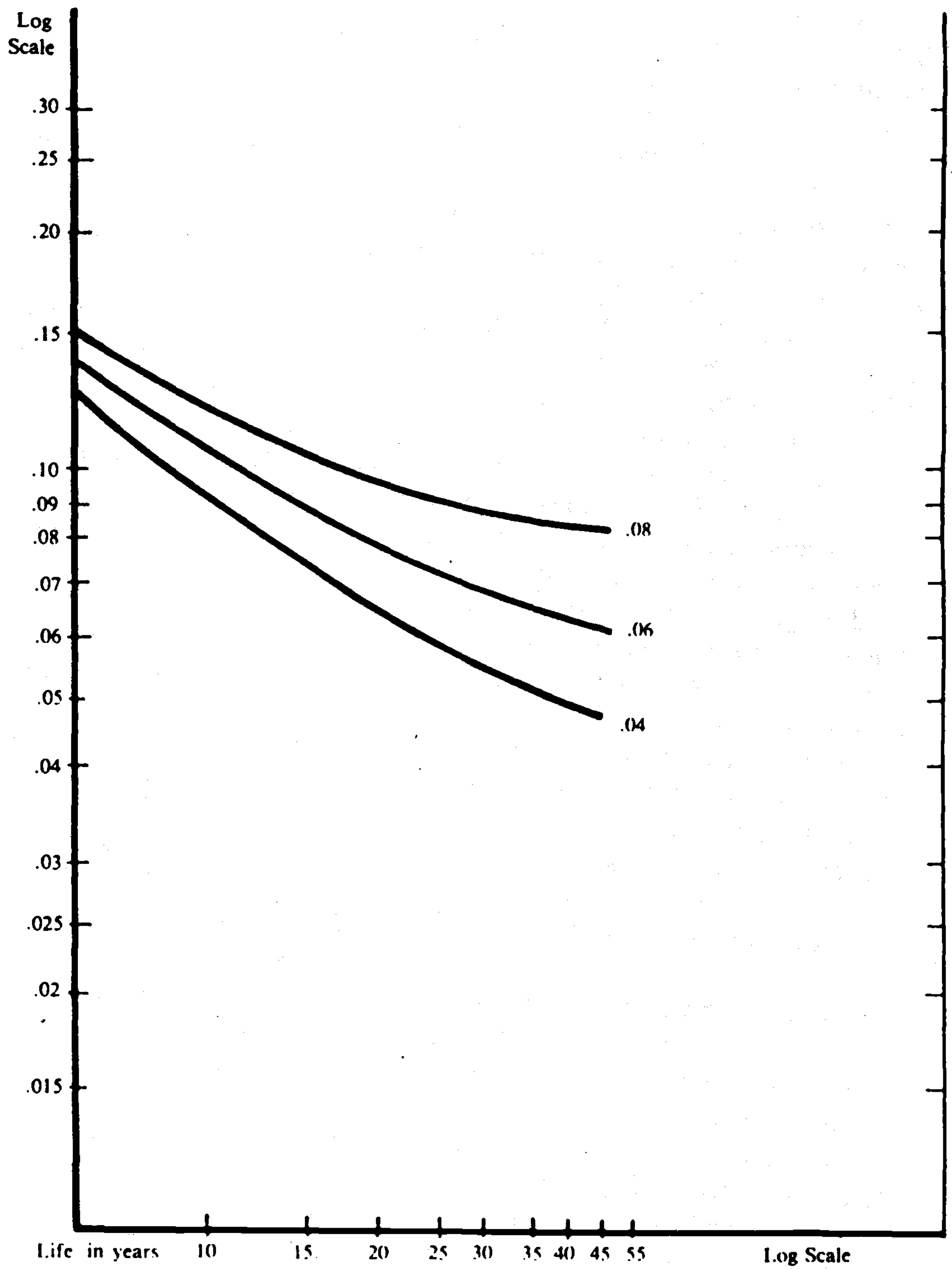
This sharply limits the advantage a CIA may claim by virtue of being a defender. Not all its capital costs are sunk, but only some fraction, depending on the actual pattern of uneven aging.

This economic liability of staggered CIA's may be a political asset, however. Owners of nearly new apparatuses view it as the greatest hardship and oppression to be assessed for a mass system that duplicates their apparatuses' function. A loss unequally distributed arouses more effective political action than the same loss spread evenly, so the plight of the few is overweighted.

The problem is aggravated a step further if demand is growing, and it usually is, because demand growth is what brings us to the threshold of the mass system; thus more than $1/L$ per cent of the apparatuses is likely to be new, and more than $1/L$ per cent is likely to require capital costs next year. This is because there is new growth on top of replacement.

FIGURE 3
CAPITAL RECOVERY FACTORS (CRF) FOR SELECTED LIVES AND INTEREST RATES

$$CRF = \frac{i}{1-(1+i)^{-L}}$$



b. Marginal cost pricing and plant decay. As the marginalists have shown, after a plant is built, price should be set to equal short-run marginal cost (*SMC*). Long-run marginal cost (*LMC*) becomes irrelevant. Economists have developed this most in respect to demand shifts: peak load pricing responsive to daily, weekly, and seasonal cycles of demand. They have done less with demand growth during plant life (like Hotelling's bridge that gets crowded). First, there is another sort of shift in optimal price which is critical and general and which the literature neglects. This is from the slow rise of *SMC* that occurs over life. Like people, old pipes suffer from arteriosclerosis, failing joints, and fatigue: in short, they "decay."

As a plant decays, it loses capacity. It becomes in effect a smaller plant. The *SMC* curve rises; optimal user charge rises; optimal volume falls; and net benefits fall.

Anticipated short-run behaviour bears on long-run decisions. Perhaps the prospect of decay and our response to it should modify our criterion of when to build a plant to begin with. It certainly requires us to specify what we mean by *LMC*. This is a useful exercise in any event. Many find *LMC* an elusive concept, even in the absence of decay—some, indeed, deny it can be defined usefully at all. In the process we can develop some background theory needed to handle the tricky problems concerning ripening, treated next, that arise when demand grows year by year. We need insight into the relations of *SMC* and *LMC*. We need some capital theory to allocate capital costs over years. This is central to specifying a replacement date (R-date) because the date may vary with how much capital cost we allocate to year zero.

Micro and capital theory have each been able to skirt these questions thanks to two different plausible devices which are good enough for rough purposes but not for facing the matter of plant decay.

Standard micro texts generally note that at the ideal plant size, when *SAC* tangent *LAC*, it is also necessarily true that *SMC* = *LMC*. They are silent on plant decay (and usually on demand growth, too), diverting us with the old tale of Viner's draftsman, so there is no problem. Short-run and long-run marginal cost pricing are the same. This is indeed a useful and central relationship, but the device needs more work to guide us when short-run average costs rise each year. What do we do when *SAC* decline to stay put?

The device of capital theory is level yearly amortization. By this technique, capital charges are spread evenly over life, so *FC* are doubly "fixed"—that is, invariant with time as well as with volume, although in production economics the meaning of "fixed" is only the latter.

Now level yearly amortization happens to be theoretically correct in the special "one-hoss shay" case when a plant does not decay—$S\acute{A}C$ are fixed—until it suddenly collapses (see Appendix 1). This is not because depreciation is constant over life but because the sum of interest and depreciation is. Value depreciates, even though service flow does not decay, because collapse nears and future life declines with each passing year.

This device lets us follow a plant through life and never change the fixed costs (*FC*) or the variable costs (*VC*). Both are invariant. And it helps us forget that life itself is a variable, subject to economic constraints, and that economists should not just accept given lives but help determine them.

Potential problems arise because the use of level amortization is so seductively neat a solution to a forbidding problem that we are tempted to use it always, even though, in general, plants do decay. This could lead to some error in deciding when to build a plant by having us allocate too little fixed costs to year zero and be overenchanted by the low variable costs, which are minimal in year zero.

Now if demand never grew over time, this error would be easily avoided by requiring candidate plants to prove their worth over whole life as well as in year zero. A plant would be submarginal unless the present value of the sum of *DMC* over life fell below the present value of the sum of *DMB*. We could convert each of these present values from lump sums into the annuities of the same present values, call them the *DMC* and *DMB*, and proceed as before.[8] (See Appendix 1.) Allocation of *FC* among years would not be necessary.

The problem arises when demand is growing steadily, year by year. This is a condition we have assumed away but will consider from here on. Now year zero must be evaluated singly. When demand is growing there are many plants that will be supramarginal over whole life, some of which will be submarginal in their early year—"premarginal." To select the best first year, which is the R-date we seek, it is helpful to screen out years that are premarginal. For this we must allocate *FC* among years. Benefits in year zero must cover not just the low variable costs but also the fixed costs of year zero. There is nothing that fixed in fixed costs that prevents our withholding them, before we invest them, from any first year that does not cover them. Capital costs are indivisible over life, but that does not prevent our advancing or retarding the whole lifespan.

If we believed, as many people do, that amortization schedules are just arbitrary, we would be tempted to build as soon as *DMB* covered *DVC* alone—no need to cover arbitrary costs that are made up later. However, common sense says we should at least earn interest on capital (C_0). Committing C_0 an extra year is clearly a cost. Even if the date of plant demolition were preordained, so that an extra first year appeared to come free of extra C_o, it would tie up C_o an extra year.

It is on the subject of depreciation that the issues arise. Under level amortization, depreciation is very small in year zero and rises each year by the interest rate (i) (see Appendix 1). This is based on the assumption of no decay and is materially slower at first than true depreciation.

There is no market in used mass systems, so what is the value that depreciates? The plant is valued internally. Value is the present value of future service flow (SF). True depreciation is the drop in this present value each year.

For decaying plants, the loss of year zero lowers value more than for one-hoss shays, because this is the best year of their lives. This is especially strong when demand is growing. There is a real benefit in matching the best plant years with larger demand; hence a real loss in wasting the best years on the small, early demand. In the extreme, you can imagine a marginal case where demand grew slowly at the precise rate a plant decayed, so that by building a year too early every year was changed from net gain to net loss.

In the mathematics of depreciation it turns out that the lower VC of year zero of a decaying plant are exactly offset by higher depreciation (see Appendix 1). Essentially this is because depreciation is the loss of the service flow of the first year (SF) less the growth in present value (PV) of SF_2 to SF_L, all of which appreciate by i because they move a year near the present. But this growth is the same as interest on remaining capital value because

$$\sum_{2}^{L}\left[SF_n (1 + i)^{-n}\right]$$

is remaining capital value. Since depreciation $= SF - i \cdot PV$, SF $=$ depreciation $+ i \cdot PV$. Thus true depreciation plus interest always exhaust SF. FC and SF are equal. Lower VC mean higher FC of like amount.

A common sense explanation is that you do not want your plant suffering decay when few people are using it. The decay of each year is an input given to the customers of that year, and their demand should cover it. Whatever its fragrance, it would be wasted on the desert air. This incidentally points up what should be obvious, that my "decay" refers only to time depreciation. Use depreciation is a VC, and should be plugged in there.

The upshot is that a plant of given C_o and PV should not be built any earlier by virtue of lower VC (hence higher SF) in its first years. A marginal candidate decaying plant whose C_O equals *PV* from future *SF* does not yield any first year surplus by virtue of lower early *VC*. The total costs allocated to year zero are independent of the actual time patterns of decay, not because annual amortization is level, but because it is not: it compensates automatically for higher yearly *SF*, which are implied by lower yearly *VC*.

So there is nothing about plant decay to make us build earlier than we would if the same present values were derived from fixed short-run cost curves. This may give us more confidence in the use of the simplifying concept of *LMC*. Micro theory does not stray by neglecting decay; capital theory may by clinging to its one-hoss shay.

There is another related factor, however, which does let us build a little earlier. That is locational obsolescence (*LO*). *SF* declines not just because of physical decay but also because of *LO*. *LO* is the yearly rise in the opportunity cost of the site under the plant, equal to interest on current site value (S_n).

Growing demand raises gross *SF* each year, but it raises $S \bullet i$ by more and therefore reduces net *SF*. As life progresses, this hastens death and speeds the drop in plant value. Now this might seem just like decay, but it is different. *LO* is independent of when a plant is built. *LO* reflects that a plant becomes too small, as demand grows, for the site it preempts. If a plant of given size is built a year earlier, this does not advance *LO* by a year (assuming demand exogenous).

So the *LO* of the first year is free. The effect of *LO* on end of life depends on the size of and demand on the plant, not the year it is built.

Thus, in assigning *FC* to year zero we must distinguish decay from obsolescence. Decay is the drop in the part of *SF* imputable to the plant capital, and *SF* imputable to plant capital is gross yield above the rising base, $S \bullet i$. The drop in value due to obsolescence is not a cost assignable to year zero.That is, take any R-date and ask, "Why not build a year earlier?" It would advance decay, that is a cost. It would advance the interest bill, that is a cost. But it would not advance locational obsolescence at all. *LO* derives from the rise of *S*, which is exogenous, except as hereafter noted. (If demand growth results from building the plant this turns out on balance to argue for earlier R-date, anyway.)

Introducing *LO* leads us into a whole new world of issues that bear on replacement policy. These are so important, and modify our findings so much, that they deserve separate treatment, which follows.

GROWING DEMAND

The analysis up to this point has been one of comparative statics and therefore incomplete and a little unworldly. We have determined how far to the right a demand curve must be to warrant building the mass system. This follows the simplified conventions of micro theory. But implicit in the question is that the demand curve is travelling, inching continually to the right, year by year. This is a more general condition, certainly. It requires we adapt the simple solution to a dynamic world.

a. The rise of attainable unit costs. With steadily growing demand, it now becomes impossible to attain unit costs as low as shown in the conventional *LAC* envelope curve because demand passes through the optimal intersection of *SMC* and *LMC*, where *SAC* tangent *LAC*, only once in any plant's life range.[9] During the rest of a plant's life it is the wrong size, so unit costs lie above the envelope curve.

In addition, the entire *SAC* curve rises because of shorter life. That is, initial capital costs (C_o) can be spread over fewer years. Locational obsolescence sets in as soon as a plant is built because growing demand keeps raising potential rent from a higher future use for the site—in this case, the intersite. This higher future rent from a second generation (G_2) challenger will shorten life and raise annual capital costs. If abrupt and rapid enough it can render a good plant submarginal.

This is no sort of tragedy, to be sure. Higher demand means higher rents, and it is only to take full advantage of these new opportunities that we raise annual costs. The only cause for dismay is that we have more thinking ahead to do to adapt to the new condition. We still want to optimize, and the first viable plant on Figure 2 may be submarginal if it will not last long enough to pay out.

Conventional analysis is silent on plant life and its determinants, but by silence implies there will be, ultimately, replacement by an identical plant yielding the same rent. The need to replace results only from physical depreciation. When we add locational obsolescence, of course the extra factor shortens life.

Coupling growing demand with the decay factor treated earlier, we see a double-barrelled effect on life. As a plant ages it loses capacity; yet growing demand calls for greater capacity. The combination may lead to materially shorter plant lives than do static assumptions.

On the other hand, life may shorten a good deal without raising yearly *FC* very much. Putting it another way, it is possible to recover initial capital (C_o) many years sooner by only a small increase in annual charges. This is because the early payments are mostly interest (see Appendix). Table 3 shows the capital recovery factor (*CRF*) for different plant lives and at different interest rates. (*CRF*) recall, is the annual instalment needed to retire a debt of 1 in L years at i interest.)

TABLE 3

CAPITAL RECOVERY FACTORS (CRF)
FOR SELECTED LIVES AND INTEREST RATES

$$CRF = \frac{i}{1-(1+i)^{-L}}$$

Interest rate	Life (years)	10	15	20	25	30	35	40	45	50
.04		0.123	0.090	0.074	0.064	0.058	0.055	0.051	0.048	0.047
.06		0.136	0.103	0.087	0.078	0.073	0.069	0.066	0.065	0.063
.08		0.149	0.117	0.102	0.094	0.089	0.086	0.084	0.083	0.082

$$CRF = \frac{i}{1-(1+i)^{-L}}$$

Note that you can always cut life in half without doubling the annual charge, even in the upper left reaches of Table 3 which represent short lives and low interest rates. In the more relevant lower right, you can halve life from fifty years to twenty-five years by raising the instalment from 0.082 to 0.094, that is by 15 per cent. Even this overstates the true elasticity involved because of the gross change assumed. If we cut life from fifty years to forty-nine (by 2 per cent), *CRF* rises from 0.81743 to 0.81886 (by 0.17 of 1 per cent). (Note in passing that *CRF*'s are much more sensitive to interest rates.)

Figure 3 shows *CRF* as a function of life on a log-log scale, where negative slope is elasticity. The curves are concave upwards and flatten out on the right, approaching the interest rate as an asymptote.

In my experience, these functions and their implications are not widely understood. It is generally regarded as prodigal to think of amortizing a major public work over a small number of years, such as fifteen or twenty. And yet economies of longevity are so largely exhausted by spreading a cost over twenty years that further savings are trivial compared to other common factors. Looking at the converse, that means the extra yearly overhead cost due to shortening life is small relative to the gains from greater adaptability to growing demand.

We need not raise the *SAC* curves much on account of shorter lives until we get into short ones under fifteen or twenty years. However, that is exactly the range of lives we may expect to get into under conditions of progressive obsolescence. Since a plant becomes the wrong size quickly after being built, we would build a new one every year, fitted to current demand; if there were no economies of longevity at all. Thus we could stay on the envelope *LAC* curve. There must be substantial extra annual costs from shortening life by another year to prevent our doing it. We will find an equilibrium life where the *SAC* curves and their envelope are all somewhat higher than under static conditions, but attainable costs are somewhat lower because we can stick closer to the higher envelope curve. I will return to this and discuss what to maximize below.

Naturally we try to minimize the rise of attainable unit costs in many ways. Shortening life is one way, as we have just seen. It raises each *SAC* curve, yet this turns out to be a small price for the great benefit of enlarging the plant more often to fit the shifting demand.

Other adjustments are fairly straightforward. Given that replacement must be discrete and periodic, a wide range of each *SAC* curve becomes relevant rather than merely the single tangency with *LAC*. A plant may be adapted to a wider range of demand in a variety of ways. For example, small rise of unit cost at the optimal volume may be absorbed in order to lower the *SAC* for other volumes. Thus, added outlays for tighter joints will allow later use of higher pressure to increase flow in pipes of fixed diameter.

As will become clear, only volumes larger than optimal are very relevant. This adaptation thus becomes similar simply to building a larger plant, analyzed below.

An obvious adaptation is to substitute variable for fixed costs. Greater use of pressure boosters in preference to line material is an example.

Another adaptation is to cut down on costs whose function is mainly to lengthen plant life. Longevity is now limited by capacity rather than by durability. Outlays to extend physical life beyond economic life are wasted. Thus the use of costlier or bulkier materials whose virtue is greater resistance to corrosion and decay after thirty years is unnecessary. The net result of these savings does not cut annual costs from the static level, because if outlays for durability were well advised under static demand they would lower annual costs. The savings here merely temper the rise of annual costs.

All that has its effect on the differential advantage of larger plants. Growing demand raises attainable unit costs more for smaller than for larger plants. This is partly because the *SAC* curves become flatter as plants become larger, so demand curves can travel further without raising *SAC* so much. They become flatter because at larger capacity any absolute increase of load is a smaller percentage of the whole and presses less importunately on plant capacity. Second, larger plants come nearer to exhausting all net scale economies. This reduces the differential advantage of the second generation plant (G_2) over its predecessor (G_1), so the life of a larger plant of G_1 is less shortened by its G_2 challenger than is the life of a smaller plant. This can be a paramount factor for the smallest challenger plants that appear viable on the static *LAC* curve. Their yearly *FC* rise so high they become submarginal, even though growing demand increases their gross yearly benefits in later years.

Also, a lower *AFC* for larger plants means less absolute extra *AFC* of such shortening as is necessary, the extra being a proportion of the base. Note that the share of *AFC* in *ATC* drops as we move right on the curve to larger plants. That is, larger plants spread their overhead among more buyers, but do not much reduce variable costs per buyer, a relation consistent with observation and theory.[10] So the percentage decline of *AFC* is greater than that of *ATC* for larger plants. Since it is the yearly *FC* which rise when we compress life, and growing demand requires us to compress life, this is an added advantage for larger plants, beyond what shows on the static *LAC* curve. Third, the optimal time to begin life of a larger plant turns out to be earlier than the tangency of *SAC* and the envelope—a matter explored below. This lets it lead demand by some years, spreading cost over a longer life.

The combined effect of these cost changes is to make us think more seriously about deferring R-date so we may build a larger plant. However, we have not entirely settled how demand growth affects R-date anyway. In the process of doing that we can weave all factors into an integrated solution.

b. Ripening: when to build what. The time to build any given plant (P_n), we have seen, is the first year when current benefits cover variable costs plus the fixed costs correctly assignable to year zero, these covering interest, depreciation, and decay, but not locational obsolescence.

To build sooner would lose net benefits in the first year. Benefits might cover variable costs, but they need also cover fixed costs in year one; that is, they must yield interest to warrant committing capital during that year and to warrant losing some of the capital by depreciation. There is in general no compensating gain from plant one (P_1) to requite any loss taken in premarginal years. The positive rents in later years would be reaped just as well by a plant built later.

This does not mean we altogether abandon the behest to build ahead of demand to achieve scale and lower costs. The criterion still has us build a plant when demand is too small to achieve minimum *SAC* when built, just as in the static case. *SAC* then declines as demand grows until the minimum. In addition, we will see there is some gain in advancing the second generation plant by a year, but this will fit into place in the later discussion of future generations. There may also be a stimulus to demand, discussed later: here we assume demand growth is given exogenously.

To build this plant later would lose a year's rent, again for no later gain. So the optimal renewal date (R-date) is quite determinate—for any given plant.

Let us begin with the first plant whose first year yields a positive rent over the CIA defender, and call it P_1. We rule out all prior plants and dates, even though some of them would yield positive net rents over life, because it is self evident there is no gain in losing money when you do not have to.

By that means we have already decided P_1 is superior to prior plants of positive net value. Is it not then also possible that P_1 is inferior to a later plant of still higher net value? We would not build P_1 any later; but we might wait to build P_2, a larger plant of lower SAC over most of its life as demand keeps growing. It is not enough for the challenger to prevail over the defender of the past; it must outdo other challengers from the future.

This is the old doctrine of ripening associated with R.T. Ely. "If I buy land and hold it for appropriate use, I perform social service. A lot suitable for a fine downtown office building may otherwise be improved with a very different, inferior building and hinder permanent improvement. . . ."[11] "It would be in the end a waste to put upon this land inferior buildings which would have to be torn down."[12]

Ely asked his doctrine to carry more weight than it would bear in the suburban land boom of the twenties, and he and Wehrwein finally accepted Simpson and Burton's revised metaphor of "cold storage."[13] Touring the suburbs today, one gets an urge to take out Simpson and Burton again. Yet there is something to Ely's idea, which has been around at least since Ecclesiastes III: "To everything there is a season, and a time to every purpose under the heaven." Perhaps he applied it to the wrong sites at the wrong time. It is more compelling apropos our intersite sector, where the future promises not merely a higher use but lower unit costs, too.

Since the larger plant (P_2) has higher costs at first, or requires larger volume to achieve given costs, it is still green when P_1 is ripe. If we want it we are better off to wait for it. The issue, therefore, resolves itself into defining and balancing the gains and costs of waiting. Many economists at this point advise one to select the future plant of "highest present value." It is a good reflex, yet not good enough, for after each alternative plant there looms a chain of successors. We will presently develop some algebraic tools to handle this complication in order to maximize the present value of the series (*PVS*). We will see that this manipulation is consistent with our basic criterion of maximizing rent, whereas maximizing present value of the first generation plant (G_1)

alone is not. The mathematics becomes simple, however, only by first defining terms and clarifying issues. Let us itemize and analyze the gains and costs of waiting.

The Gains of Waiting

The basic gain is the lower SAC_2 curve of a larger plant (P_2), not just at the tangency with LAC but increasingly with time and larger volumes. We have already stressed how the advantage of larger plants gains by dynamic growing of demand.

The lower SMC_2 allows larger volumes, as well, by repetition of the reasoning expressed in Figures 2 and 3.

This is some gain, but it is mixed, for the marginal volume is at a low reach of the demand curve, limiting its value. Also the future rents must be discounted and the discount factor grows geometrically with time whereas demand (by assumption) grows arithmetically and volume slower than that as SMC rise.

Gain two, a small one, is that the optimal R-date of all plants (P_n) larger than P_1 (the first marginal one) is slightly left of SAC_N tangent LAC. The difficulty of this finer point is greater than its importance. I include it to anticipate any confusion it might occasion if it were unstated.

The R-date advances because all tangencies to the right of P_1 (that is, SAC_1 tangent LAC) must yield a surplus. This follows because P_1 breaks even at its tangency. When the demand curve has risen, it lies above LMC all the way to the new intersection with determines the size of P_2, where $LMC = SMC_2$ and LAC tangent SAC_2. P_2 should therefore be started before the demand curve has risen this far; otherwise it is overripe.

Of course the in-between plant smaller than P_2 whose SAC_N tangent LAC at this earlier time is also overripe then, and so on back to P_1 where the R-dates converge. So there is no ambiguity over what plant is ripe on any given date.

Thus we see a reason why larger plants can spread their FC over more years. As demand grows the plants becoming successively ripe are larger than those whose SAC tangent LAC at the intersection of demand and LMC. And this effect grows as the lifetime rents of plants grow.

Gain three is better foresight. As a generality this argument has no value, for it will be equally true next year—more so, perhaps, because next year's optimal plant requires a longer forecast. There is only a gain if the current rate of clarification is above the future rate of improved knowledge of demand and costs.

Claiming "uncertainty" is not a good stall, either. It is valid if you suspect future demand will grow faster than forecasted. On the other hand, if you suspect it will grow more slowly, that argues for building earlier—the small plant, of course. If you think the fog will lift in fifteen to twenty years, you can, in the meantime, amortize a small interim plant.

Gain four is added productivity. This is a normal expectation. It means a later plant enjoys lower costs and gives better service, thank to technological advances as well as more appropriate scale, adding to the advantages of larger plants.

Gain five is higher costs of the existing defender CIA. It may seem odd to call this a gain, but the rents of our alternative challenger plants are just their advantages over the CIA. The worse the CIA, the greater the gains. It satisfies common sense to think that one of the gains is another year's use from the old CIA plants.

Gain six can be reduced cost of inputs. This might seem far-fetched today, but it is not. When we correct for depreciation of the dollar, we must take into account the fact that the real cost of most raw materials has fallen for some time and may continue to. The most important input by all odds is borrowed money, and there is always a chance of lower future interest rates, although I myself do not foresee that. It would be wrong to argue that accelerated future inflation is equivalent to lower future real rates of interest, because money borrowed this year for a long term will enjoy this benefit almost as much as money borrowed in the future.

Gain seven is that a federal programme may help pay. Socially this is spurious, of course. It is an effective delayer of many good works, however, and socially quite damaging, especially in conjunction with other federal programmes of the reverse twist where works are accelerated to capture matching funds before they run out.

There, then, are the wiles of Scheherazade. One should hear them sceptically, for with these a confirmed temporizer can procrastinate without end. So let us now consider the other side of the waiting game.

The Costs of Waiting

The major loss is deferral of rent or of net benefits. Costs, to be sure, are also deferred, but waiting defers benefits over costs, and there is the loss.

The lost rent of P_1's year 0, taken singly, is of little moment, for year zero is marginal. But by the time P_2's first year is marginal P_1 would have been postmarginal and rent yielding. Looking forward and comparing the respective successor plants in the P_1 series, we see the same is true *a fortiori* because of P_2's longer life; but I defer this for separate treatment.

On balance, then P_1 yields smaller net rents than P_2 does but yields them sooner. Sacrificing P_1 for P_2 is the same as making an investment. The thing invested is the sacrificed net present value of P_1. The thing gained in return is the net present value of P_2 later when P_2 is ripe. The cost is interest over the period of waiting on the value invested. The investment is good if it yields a competitive rate of return. Let MPV_n be the moving present value at the start of year n for plant n; let i be interest rate. Deferral to m pays if $MPV_m \geq MPV_n(1 + i)^{m-n}$. This is equivalent to the rule of maximizing present value (from a fixed date of reference) and is all right as far as it goes, which is only into the first generation. We will see that it becomes complete when extended to comprehend all.

Cost two is deferment of demand. I have treated so far only the exogenous aspect of demand, but where the mass system is prerequisite to immigration there must be some lag in attracting demand. It is not enough to proclaim an intent to build. To pour cement is more credible, and the market responds. The present buildup of future demand, discounted, may be treated as a current revenue. While that may sound fuzzy, it has an observable index in the rise of land values before use. This is too controversial and complex a topic to elaborate upon here, but there is something to it. The Haig-Simons doctrine tells us that capital gains are income at the time they accrue. If so, advanced land values represent benefits at the time they advance. And we must not fail to treat the enhanced capital value of the intersite enterprise and the advanced treasury equity in the same way. On the other hand, we must beware of soft land values traded on thin equities with easy credit. We must beware of land bubbles. We must beware of claiming credit for exogenous demand increments. And we must beware of purely redistributive gains.

Cost three is future inflation of input prices. "Build before costs rise again": one is given this rationale constantly. It is not generally valid, because general inflation is an illusory cost of waiting. But financing with long-term bonds changes that. Future inflation means a lower real rate of interest than the nominal one. Lower interest rates reduce the disadvantage of incurring costs sooner rather than later.

However, this is double-edged. Lower interest rates also reduce the disadvantage of receiving revenues later. Since revenues are greater than costs, and larger plants mature later than smaller ones, the revenue side is stronger. The net effect of lower interest rates is to favour waiting, as summed up under cost one. Cost three turns out to be gain eight.

Cost four is that demand may not grow as expected. This cost is very small, because the lost years of P_1 were close to marginal and the option has not been lost of building P_1 later.

Cost five is the loss of limited federal matching funds, discussed under gain seven. It is spurious, from the social point of view.

Cost six is the loss of or failure to acquire the site by franchise, licence, monopoly, or the like. This is generally purely acquisitive and therefore spurious, although occasionally an easement might be saved from irreversible loss to a lower use. This is a paramount motive in practice for premature extensions of mass systems into unripe territory. In conjunction with the decentralist bias in rate making it makes mischief too extensive for a few words to damn adequately.

The Cost of Deferring Succession

All the gains and costs of waiting can find sockets now in a simple model where the basic cost is interest on the moving present value (MPV) and the gain is growth of MPV.

Finally we must ask, the present value of what? Many economists are satisfied to look only at the first generation, believing later values to be negligibly small. In the present situation of growing demand, decreasing costs, and rapid obsolescence that cannot be the case. We have seen that the life of the first generation in any series (G_1) is compressed by the rising challenge of G_2, so much so as to raise yearly *FC* materially. We have seen this implies a life compressed to twenty years or so. Values after twenty years, especially higher ones, are not negligible today. If they were, they would lack power to make us compress life and raise yearly *FC* in the present. Sacrificing these successors of P_1 must add materially to the cost of sacrificing $P_1 G_1$. At the same time gaining the successors of P_2 adds to the gain of waiting, so it is not altogether one-sided, but a question of which is larger.

As it turns out, a showdown based on G_1 alone is badly weighted against the smaller plant, P_1. With growing demand, naturally P_2 shines because it is geared to growth. Comparing only the G_1's of P_1 and P_2, the analyst may either assume equal lives or let P_2 last longer. If he assumes equal lives he stretches P_1 into a Procrustean postretirement career beyond its capacity where it compares badly with the larger P_2. Even worse is to let P_2 last longer: *PV* varies directly with life, so the *PV* of P_2 gains from the extra years denied to P_1.

Less biased is to find present values by capitalizing yearly rent from P_1 and P_2. Yearly rent is normalized to adjust for different lives. It is a sort of yearly average of net benefits over life, adjusted for time (see Appendix 2). However, this turns out to be a sneaky way of introducing successor generations: "capitalizing" assumes perpetual life. Capitalizing yearly rent from G_1 gives the *PV* of G_1 plus an endless chain of identical successors (see Appendix 2).

Capitalizing, then, is an improvement, but it gives the *PV* we seek only in the special case of identical successors. It falls well below the *PV* we seek now, that of a series of ever grander successors that expand ahead. In the special case of constant rent, the cost of waiting, $MPV \cdot i$, is simply yearly rent, because $MPV = \text{rent}/i$. In the present case of higher rent from superior successors, the cost of waiting is more than the lost rent of the year's deferral. There is an extra cost due to deferring higher future rents. The extra cost may also take the form of shortening G_1, but optimal life is determined so that this cost would equal the other. That is how one chooses the life that maximizes PV for any series.

Let us lay a basis for estimating how much of waiting cost stems from later generations. Let G_n be the present value of G_n on its R-date. S is the present value of the site. Begin with the case of equal successors so that every G_n is the same :

$$S = \frac{G}{1 - (1 + i)^{-L}}, \qquad \text{(see Appendix 2).}$$

The share of S derived from later generations is:

$$\frac{S - G}{S} = 1 - \frac{G}{S} = (1 + i)^{-L}.$$

Any value of $(1 + i)^{-L}$ is easy to look up or to estimate from the quite accurate rule of thumb that the value halves every $0.72/i$ years. At 7.2 per cent and twenty years it halves twice; thus the share of future generations is 1/4. That means they add 1/3 to the PV of G_1 and 1/3 to the cost of waiting.

That is a minimum estimate based on equal successors. With superior successors S rises, so the ratio G/S falls and the extra cost of deferring future generations rises. In a simple growth model:

$$S = \frac{G_1}{1 - (1 + i)^{-L}(1 + g)},$$

where g is a constant percentage by which succeeding generations outvalue one another (see Appendix 2). Now the share of future generations is:

$$1 - \frac{G}{S} = 1 - \left[1 - (1 + i)^{-L}(1 + g)\right] = (1 + i)^{-L}(1 + g)$$

If $g = 100$ per cent, $(1 + g) = 2$ and the share of future generations is 1/2, meaning the PV of the site is doubled.[14]

That model needs modifying because the prospect of rent's doubling every twenty years forever is hard to credit. Eternally unfolding vistas are all very well in the afterlife, but on earth that could cause a lot of trouble, especially in this case, where higher density is the major factor increasing rent. Fortunately, most of the effect comes from the first doubling. Suppose that $G_2 = 2G$ and that all future $G_{>2} = G_2$, that is, remain on the new plateau. Then the share of $G_{>1}$ is:

$$1 - \frac{G}{S} = 1 - \frac{1}{1 + \dfrac{2}{(1+i)^L - 1}} = \frac{2}{(1+i)^L + 1} = \frac{2}{5}.$$

The share of 2/5 means future generations add 2/3 to PV_1 of G_1. This last model is in general the most relevant to our case, where there are great gains at first which taper off toward a limit.

In sum, the cost of waiting is the lost interest on G_1 plus the deferral of higher later rents. The latter is interest on the PV of later generations. It adds as much as 100 per cent to the cost of waiting.

Now let us see if there are any offsetting gains from waiting, derived from future $G_{>1}$ in the series following P_2. There are some gains whenever any G_n of P_2 yields more rent than G_n of P_1. Some such gains are plausible by repeating the same reasoning that makes P_2 worth more than P_1. The G_2 following P_2 is built later than the G_2 following P_1 and may enjoy greater demand. Indeed, if the rate of advantage is so great that the PV_0 of P_2G_2 equals PV_0 of P_1G_2 (and so for all G_n), then the gains of waiting equal the costs. In order for this to be, the MPV of G_2 must grow at the rate of interest.

In general,that is unlikely. First, P_2 has a longer life than does P_1. The deferral of G_2 is the sum of the wait plus the extra life. Thus to secure the benefits of P_2 we doubly defer the greater benefits of G_2.

Second, added gains taper off. It is entirely possible that P_2G_2 will be not at all better than P_1G_2, even though it is built ten to twenty years later. That is the model I suggested above as most relevant.

On the other hand, there are circumstances where waiting for P_2 will itself raise the value of G_2. Suppose we are subdividing land in conjunction with G_1, and every year we wait lets us raise density by reducing lot size. Subdivision stamps a density pattern on land that may endure. Higher density raises the value of G_2 as much as G_1. To avoid the irreversible damage of permanent oversized lots is worth some waiting.

Any preemptive lower use that we let take root may be visited on the sons, and the sons of sons, even unto G_7. There is always great delight in revisiting Sam Walter Foss's "The Calf-Path."[15]

Foss's point, however, was not against the calf, but the later generations who followed him. A good deal of so-called irreversible damage is only as permanent as acquiescence lets it be, and we need not visit our sins on our sons if we empower them to remake their world for their times. Indeed, the only sin is binding future generations to former virtues.

Therefore, how much we may expect to raise P_2G_2 by waiting today depends on how inflexible a set of land use institutions we leave posterity. Granting freedom to the future frees us to raise our welfare in the present, and who can doubt that the future will benefit, too? So on balance I recommend we see the results on future generations of waiting primarily as deferral without improvement. That means all costs and no benefits. The added cost may often be, as shown, on the order of 50 per cent to 100 per cent. This fact weighs heavily for advancing R-date to today.

CONCLUSION

As always, limited space forces one to end before satisfying every expository need or treating every issue. Even if everything advanced here is perfectly correct, there is more to do. More is needed on how locational obsolescence affects optimal life. The effect of the property tax needs treatment: how does it slow private investment responsive and complementary to mass systems; how may it be used to speed response; how should ripening lands outside mass systems be assessed and taxed? May land value increments be used to measure present values of future derived demands? Answers to these questions must wait on future research. Meantime I hope these generalized theoretical notes may help those many economists who are busy proving that economic analysis applied to specifics has a powerful potential for raising welfare and can be the most relevant of disciplines.

Appendix 1

Some Basic Mathematics of Depreciation and Annualization

An asset yielding a level annual flow of a over L years has a value (V) in any year (n):

$$V_n = a\frac{1-(1+i)^{n-L}}{i} \tag{1}$$

Depreciation is the negative change in V with respect to n:

$$\frac{dV}{dn} = \frac{a}{i}\left[0 - ln(1+i)\cdot(1+i)^{n-L}\right] \simeq -a(1+i)^{n-L} \tag{2}$$

The above result is simplified because $i \simeq ln(1+i)$ and they cancel. The latter is the continuous rate of interest, conventionally designated by the Greek $ro(\rho)$. The use of this approximation does not imply any approximation in the result but a trivial adjustment for the fact that the derivative of the numerator is instantaneous, while i in the denominator is premised on annual compounding. If we were to abandon calculus and compute the growth of the numerator by simple algebra we would find

$$\frac{\Delta V}{\Delta n} = \frac{a}{i}\left[0 - i(1+i)^{n-L}\right] = -a(1+i)^{n-L}$$

with no approximation at all.

Interest on the undepreciated balance is:

$$V_n \cdot i = a\left[1-(1+i)^{n-L}\right]. \tag{3}$$

Interest plus the positive value of depreciation is therefore always a:

$$V_n \cdot i + \frac{dV}{dn} = a\left[1-(1+i)^{n-L}\right] + a(1+i)^{n-L} = a \tag{4}$$

Thus annual FC, the sum of interest and depreciation, is constant over life for a plant of constant service flow. Note that the proportions of interest and depreciation are not fixed, however. Interest dominates the early years; depreciation, the later years.

The level flow that returns a dollar of capital with interest over L years is known as the capital recovery factor (CRF), or the "annuity whose present value is one." It is the reciprocal of the present value formula when $n=0$, that is at birth:

$$CRF = \frac{i}{1-(1+i)^{-L}} \cdot \tag{5}$$

$CRF \times V_0 = a$. Thus the CRF is a coefficient that converts a lump sum present value into the level income stream from which it was or might have been derived by discounting. It thus tells you what level income stream is needed to return any capital outlay, with interest. Tables of these coefficients for different values of L and i are standard, being the basis of instalment payments to service and to retire debt.

A depreciation schedule corresponding to the level flow assumption is not, obviously, itself level or straight line. Depreciation is small at first and rises nearly to a at last. The schedule is implicit in the "sinking fund" method. Under this method, a level annuity (of an amount equal to a discounted over L years) is paid into a fund. The fund grows at compound interest. At maturity the fund equals the original cost. The annuity (b) is:

$$b = V_0 \frac{i}{(1+i)^L - 1} \left(= a(1+i)^{-L} \right) \tag{6}$$

The fund (F) at any time has the value:

$$F_n = b \, \frac{(1+i)^n - 1}{i} \; . \tag{7}$$

Depreciation is the growth of the fund:

$$\frac{dF}{dn} = \frac{b}{i} \ln(1+i) \cdot (1+i)^n \simeq b(1+i)^n. \tag{8}$$

To show that the growth of the fund is the same as the depreciation of the asset, we need to show that

$$\frac{dF}{dn} = -\frac{dV}{dn}$$

which, we will recall, is $a(1+i)^{n-L}$.

$$\frac{dF}{dn} = b(1+i)^n = a\,\frac{1-(1+i)^{-L}}{i}\;\frac{i}{(1+i)^L-1}\,(1+i)^n$$

$$= a\,\frac{\left[(1+i)^L - 1\right](1+i)^{-L}}{(1+i)^L-1}\,(1+i)^n = a(1+i)^{n-L} = \frac{-dV}{dn} \; . \tag{9}$$

Alternatively, this result also follows from an earlier definition of b:

$$b = a(1 + i)^{-L} \tag{6}$$

$$\therefore \quad b(1 + i)^n = a(1 + i)^{n-L} . \tag{10}$$

A plant that yields a declining annual flow "decays." This increases its depreciation in early years, in relation to later ones. It remains true, however, that the annual flow of every year equals the sum of depreciation plus interest. This is most easily seen by conceiving of depreciation as the loss of the value of the first year's flow less the appreciation of future years' flows which all move one year nearer the present. This appreciation is their present value times i, thus equalling interest on present value of the plant. So depreciation equals flow less interest, and therefore, of course, depreciation plus interest always equals the flow.

Algebraically, this approach to depreciation is as follows:

1. For a nondecaying plant

$$\text{Depreciation} = a\left[(1 + i)^{-1} - i\left[\frac{1 - (1 + i)^{n-L}}{i} - (1 + i)^{-1}\right]\right]$$
$$= a\left[1 - 1 + (1 + i)^{n-L}\right] = a(1 + i)^{n-L}.$$

As before, depreciation plus interest = a

2. For a decaying plant

$$V_n = a\frac{1 - (1 + i + d)^{n-L}}{i + d}$$

where d is a decay factor added to interest.

Proceed as in (1).

If the two plants begin with the same flows (a), the decaying plant has the smaller value. For it to have the same value, its initial flow (c) must satisfy:

$$c\frac{1 - (1 + i + d)^{-L}}{i + d} = a\frac{1 - (1 + i)^{-L}}{i}$$

$$c = a\frac{1 - (1 + i)^{-L}}{1 - (1 + i + d)^{-L}}\,\frac{i + d}{i}$$

The higher flow of c in year zero owing to higher volume and lower variable costs is exactly balanced by higher depreciation. The extra depreciation is c-a. Interest in year zero is the same and slightly less thereafter.

Appendix 2

Some Basic Mathematics of Rent and Capitalization

Rent from an L-year cycle of investment and liquidation on a site is the level annuity whose present value equals that of the cycle. It is found by a threefold process of discounting, summing, and levelling. The exact recipe is:

$$\text{Rent} = \sum_{o}^{L}\left[(R_n - C_n)(1+i)^{-n}\right]\frac{i}{1-(1+i)^{-L}}$$

where R_n and C_n are revenues and costs dated year n, i is interest rate, and L is life.

The levelling device is the capital recovery factor (CRF), discussed in Appendix 1 and in the body of the article.

The sum in brackets is net present value, henceforth PV.

To capitalize rent in perpetuity divide by i, cancelling the numerator in the CRF. The result is PV of the site in perpetuity (PVS), assuming identical successors. The same result is obtained by summing an endless series of equal PV's separated by L years:

$$PVS = PV\left[1 + (1+i)^{-L} + (1+i)^{-2L} + \cdots + (1+i)^{-\infty}\right]$$

$$= PV\frac{1}{1-(1+i)^{-L}}$$

Maximizing rent or PVS with respect to life yields identical solutions, since rent $= i \cdot PVS$, and i is not a function of life. Thus maximizing rent is consistent with maximizing PVS. Maximizing PV alone, however, yields much too long a life. PV continues to grow with life as long as $R_n > C_n$. Rent and PVS top out long before, since their denominator grows with life as $(1+i)^{-L}$ shrinks. They top out when numerator and denominator grow at the same percentage rates.

When successors have higher PV, $PVS \cdot i >$ Rent. For example let g be a growth factor,

$$\left[(1+i)^{L} - 1\right] > g > 0.$$

$$PVS = PV\left[1 + \frac{1+g}{(1+i)^{L}} + \frac{(1+g)^2}{(1+i)^{2L}} + \cdots + \frac{(1+g)^{\infty}}{(1+i)^{\infty}}\right] = \frac{PV}{1-(1+g)(1+i)^{-L}}$$

The pressure of this higher PVS shortens each generation, as the reader may confirm by setting the life-derivative of PVS equal to zero or, less formally, by noting how g in the denominator gives leverage to its percentage growth rate. (This model overstates the case by making g independent of life. Limited space prevents presenting the more varied models needed for greater insight.)

Just as the higher *PVS* presses each of the members of this series into a shorter life, so it shortens the life of any prior defender and advances the optimal inception of the whole series.

Notes

1 Paul Downing, "Extension of Sewer Services at the Urban-Rural Fringe," *Land Economics* (1969): 103-11.

2 P.A. Stone, "Economics of Housing and Urban Development," *Journal of the Royal Statistics Society* (1959): 417-60.

3 J. Hirshleifer, J. Milliman, and J. De Haven, *Water Supply* (Chicago: University of Chicago Press, 1960), p. 103. TVA is a classic case in power.

4 DMC is also represented as the area under SMC_D from A to the cusp, plus the area under LMC_C from the cusp to B. By deleting the reach of the defender curves right of the cusp, and the challenger curves left of it, we have a unified set of cost data joining the two alternatives. The curves may be spliced at the cusp because there the total cost is equal for both. The unified MC curve now is spectacularly broken—and significantly so—but still continuous and integrable.

5 The graphical technique here has wider applications. For example, the defender might not be a CIA, but an older mass system suffering from short-run increasing costs because of inadequate capacity or decay. Again, the choice might be on virgin land where neither alternative was a defender. Then simply add fixed costs to the CIA curves and proceed as before. Finally, the whole thing might be reversed, the mass system being an old obsolescing defender, like a passenger railway, being challenged by a CIA like automobiles.

6 Alternatively, some water cooperatives in Utah and southern California issue shares proportioned to acreage and then assess the shares. This is not a common institution, however.

7 Maury Seldin, "The Role of a Market Information System in Regulating Land Use," *Assessors Journal* (1969): 28-38.

8 Note, we cannot convert the entire curves of cost and demand in this way and use their intersection to determine volume. The intersection moves each year, and we must follow it year by year. One must itemize before consolidating.

9 "Life range" means the total time between building and scrapping. In the present context I henceforth use simply "life" to mean this range of years.

10 This may not be immediately obvious. It follows from the relation that LMC gets nearer to 100 per cent of LAC as LAC flattens to a minimum; and SMC = LMC at ideal plant

output. Our Figures 1 and 2 do not show this finer point, since a constant power rule of scale economies was followed in drawing LAC in all its reaches, so on the Figures LMC = 50 per cent of LAC at all points. The relevant visual model for this point is the U-shaped LAC of conventional micro theory.

11 R.T. Ely, "Land Speculation," *Journal of Farm Economics* 2 (1920): 121-36.

12 R.T. Ely, "Outlines of Land Economics," Mimeographed paper (Ann Arbor: Institute for Research in Land Economics, 1922), p. 104.

13 H.D. Simpson and E.R. Burton, *The Valuation of Vacant Land in Suburban Areas* (Evanston: Northwestern University Press, 1931), p. 44.

14 Actually the future share may be greater yet if lives shorten to anticipate growth, raising the value of $(1 + i)^{-L}$. S varies inversely with L. For brevity's sake I omit this point, which would also require us to recognize g as a function of life and lower it for shorter lives.

15 For the benefit of the young who no longer study the classics, this recounts how an artery of commerce came to follow a winding course originally blazed by a drunken calf.