A Review of "On the Social Desirability of Urban Rail Systems"by C. Winston and V. Maheshrib[1]

Haynes Goddard, Ph.D.
Department of Economics
University of Cincinnati
May 2009 [2]

Abstract

Clifford Winston and Vikram Maheshri attempt to use benefit-cost analysis to make a definitive statement about the social desirability of urban rail transit in the United States. Their argument is deficient on several elementary analytic and statistical grounds. They underestimate total benefits, and therefore net benefits, and their failure to examine the suitability of their data and to pay attention to the usual caveats associated with benefit-cost analysis further undermines their assertions. As a result, these findings should not be used to inform either the debate or decisions about investment in urban rail systems.

Introduction.

Large infrastructural projects in transportation and water resources are often controversial. Water resources projects are frequently alleged to cause major environmental damage that outweighs their economic benefits (think of the Florida Everglades or the dams on the Columbia River in the State of Washington). In transportation, urban highway projects are advocated by some as a method to reduce road congestion, and opposed by others because they tend to induce additional traffic and sprawl. It would seem that a careful and disinterested benefit-cost analysis of proposed large projects would always be in the public interest. But the fact is that large projects and the associated expenditure create many economic interests, especially among potential suppliers, and this fact consequently creates complications and barriers to disinterested analysis. And of course, really large projects are usually funded with government taxes, so local beneficiaries receive project service flows that are typically heavily subsidized, with the result that they become another important interest group in the projects. The infamous Alaskan "Bridge to Nowhere" is a recent

---

[1] Journal of Urban Economics, 62 (2007) 362–382, and also at at http://web.iitd.ac.in/~tripp/delhibrts/metro/Metro/on%20the%20social%20desirability-brookings.pdf

[2] I thank Winston Harrington and Jay Warner for their helpful comments and Todd Litman for his extensive suggestions.

example.

Aside from the political complications, all too often the project analysis that is conducted is a confused mixture of economic and financial analysis, and is undertaken by engineers and not economists, almost all of whom have taken an "engineering economics" course in their training. But these courses tend to be just finance courses applied to engineering projects, with typically little real economics involved in the courses, and consequently little in the project analysis undertaken by state and regional authorities. The tip-off of this confusion lies in the nearly universal computation of benefit-cost ratios rather than the analytically correct net benefits.

Further, the studies are usually contracted out to consulting firms, and the behavioral dynamic here is that negative findings may very well mean no future contracts, so these firms have a powerful incentive not to find that such projects make no "economic" sense. "Economic sense" can have at least two meanings: a) benefits exceed costs, or b) the project is cost-effective. So it is not surprising that in such studies virtually all road projects pass an economic test, virtually all water resource projects pass an economic test, and more recently, the allegation that virtually all rail transit projects pass an economic test. If anyone has had the opportunity to observe the highway analysis process up close, it quickly becomes clear that the maxim about politics applies: "if you don't like to watch sausage being made, stay out of the kitchen". It is not an intellectually honest process, fancy models notwithstanding, and it is a fair hypothesis that too many of these projects are approved that should not be, and those that are approved are too large.

Also true about such projects is that governments seldom conduct *ex post* benefit-cost analyses of projects to see if the projected returns in the original analyses were in fact realized. It is usually very difficult, if not impossible, to get an agency to fund such a study. And of course, there is self-serving bureaucratic behavior here too, because, for example, should an *ex post* study of an urban highway widening suggest that the projected congestion relief did not occur, that would cause employment problems for all the highway engineers in the state highway departments, to say nothing of commercial interests.

This situation greatly frustrates professional economists, who are trained to look as dispassionately as possible at all of the benefits and costs of a project, noting the distribution of benefits and costs by income class and other variables such as jurisdiction, i.e., who gains and who loses. But unless a project has the purpose of income redistribution, then the political beneficiaries are of no importance in the economist's analysis. The focus is on the net benefits, to whomever they may accrue, and a strong argument can be offered that this is the way it should be. With all the various interests involved, professional economists look at all large projects with skepticism: they think that many projects would not be justified on benefit-cost grounds and that there is too much private gain seeking going on, what economists call "rent seeking".

So occasionally, an economist in a university or an independent research institute will undertake such an analysis. A major difficulty such analysts confront is that they do not have funding to do a really thorough analysis, as with trip-generation models in the case of highways or mass transit, and as a consequence, the analysis has to be relatively constrained. Full blown analyses are very data and computationally intensive, and therefore expensive. So the independent economist is forced to make do with published data and perhaps some data to be cajoled out of highway and transit agencies. The models then are necessarily relatively simple as a result, running the risk that they may be too simple and may provide either erroneous or at least misleading conclusions.

That difficulty raises an important question. In the case of thorough transport project evaluation, both demand and cost modeling is very elaborate and complex, requiring hundreds of professional hours and a vast amount of data to estimate the needed relationships. These analyses, which include engineering analyses, can easily run into the millions of dollars for very large projects. So there is no way the independent analyst can rerun the modeling exercise with the *ex post* data to see if, where or why the original forecasts and estimates were correct. As a result, one cannot be sure that the findings of a relatively simple *ex post* study that either confirm or contradict previous investment decisions, do so because the project in fact was or was not justified, or that the findings are just an artifact of an incomplete *ex post* analysis. This then is the motivation for this review.

By way of some relevant background, broadly speaking, microeconomists are trained in two areas: 1) pinning down how the economic world works (mainly why people make the choices they do) through modeling and statistical analysis, and 2) using that information to assess the degree to which our public and private choices about expenditure and investment raise social welfare. Central to assessing private market functioning is the concept of "market failure" (pollution and monopoly are examples). There is a parallel concept of "government failure", which we unfortunately see in the newspapers all the time. Recent examples are the regulatory failures surrounding the mortgage crisis, or the problem of the influence of lobbyists on legislation and executive branch decisionmaking, giving new meaning to the idea that we have the best government money can buy.[3] Many observers of the decisionmaking process for public infrastructure investments conclude that there are many examples of government failure there as well – the U.S. Army Corps of Engineers and metropolitan highway planning are two of the most prominent in which project analysis is commonly poorly done or worse, intentionally distorted to favor projects.

---

[3]The intersection of the mortgage crisis and lobbying are nicely discussed in the New York Times article "Long Protected, Fannie and Freddie Ballooned", July 13, 2008. http://www.nytimes.com/2008/07/13/business/13lend.html?hp0

374 *C. Winston, V. Maheshri / Journal of Urban Economics 62 (2007) 362–382*

Table 5
Social net benefits of transit, 2000* (figures in parentheses include an exhaustive public spending cost of 10.2%)

City (agency)	Consumer surplus	Transit agency deficit	Net benefits	Congestion savings to road users	Social net benefits
New York (NYC Transit)	850	2500 (2750)	−1650 (−1900)	1195.7	−454.3 (−704.3)
Washington, DC** (METRO)	281	657 (724)	−376 (−443)	181.3	−194.7 (−261.7)
SF Bay Area (BART)	542.5	624 (688)	−81.5 (−145)	181.3	99.8 (36.3)
Chicago (CTA)	391	644 (710)	−253 (−319)	272.8	19.8 (−46.2)
Boston (MBTA)	256	701 (772)	−445 (−517)	64.4	−380.6 (−452.6)
Atlanta (MARTA)	120	424 (467)	−304 (−347)	45.5	−258.5 (−301.5)
Philadelphia (SEPTA)	54	365 (402)	−311 (−348)	77.0	−234 (−271)
Northern New Jersey (PATH)	62	141 (155)	−79 (−93.4)	6.3	−72.7 (−87.1)
Los Angeles Metro	17	477 (526)	−460 (−509)	383.8	−76.2 (−125.2)
San Diego Trolley	6.8	47.5 (52.3)	−40.7 (−45.6)	16.4	−24.3 (−29.2)
Portland, OR (TriMet)	4	213 (235)	−209 (−231)	9.1	−199.9 (−221.9)
Baltimore (MTA Maryland)	6	198 (219)	−192 (−212)	14.9	−177.1 (−197.1)
Miami-Dade Transit	Negligible***	141 (156)	−144 (−158)	16.9	−127.1 (−141.1)
San Francisco (Municipal Railway)	3	276 (305)	−273 (−301)	51.5	−221.5 (−249.5)
St. Louis (Bi-State Dev. Agency)	Negligible***	160 (176)	−159 (−176)	5.0	−154 (−171)
Southern New Jersey (PATCO)	2.57	8.77 (9.66)	−6.2 (−7.14)	Negligible	−6.2 (−7.14)
Cleveland (GCRT)	2	115 (127)	−113 (−125)	7.5	−105.5 (−117.5)
Dallas (DART)	13	443 (488)	−430 (−475)	18.2	−411.8 (−456.8)
Sacramento RT	Negligible***	96.7 (107)	−100 (−110)	4.0	−96 (−106)
San Jose (Santa Clara Co. Tr.)	Negligible***	202 (223)	−201 (−222)	11.5	−189.5 (−210.5)
Pittsburgh (PA Allegheny Co.)	Negligible***	127 (140)	−126 (−139)	3.6	−122.4 (−135.4)
Denver (RTD)	Negligible***	259 (277)	−260 (−285)	5.6	−254.4 (−279.4)
Staten Island (SIRT)	3.4	22.5 (24.8)	−19.1 (−21.3)	Negligible	−19.1 (−21.3)
Buffalo (Niagara Frontier)	Negligible***	51.2 (56.5)	−51.5 (−56.7)	Negligible	−51.5 (−56.7)
Newark (NJTransit)	Negligible***	55.1 (60.8)	−54.2 (−59.8)	1.2	−53 (−58.6)
Total			−6338 (−6984)	2573	−3842 (−4496)

* All figures are in millions of 2000 dollars.
** The actual congestion savings to drivers for Washington, DC, are unavailable, so we use the estimated savings for a comparable metropolitan area with a similar transit system (SF Bay Area).
*** Negligible consumer surplus (i.e., estimated consumer surplus is less than the error term).

Table 1. Net Social Benefits of Transit

The Larger Debate Over the Benefits and Costs of Rail Transit

There has been a continuing debate concerning the role of public transit in modern transportation systems. Critics, mostly libertarian in orientation, have typically argued that rail transit is very expensive compared to alternative technologies such as highways and that it is not cost-effective as a transport technology. Their arguments focus on costs and do not examine the benefits.[4] Proponents of rail point out that rail transit provides a wide range of benefits, including many that are indirect and non-market, and as a result are difficult to quantify, but nonetheless have to be included for an analysis to be complete.[5] Rail transit critics tend to focus on the broad transportation and land use trends that occurred during the 1950s through 2000 and that occurred in the absence of new rail investments, a period in the U.S. of massive investment in interstate highways and metropolitan expressways. Rail proponents tend to focus on specific examples of successful transit projects, or estimate future growth in transit demand and benefits.

Nonetheless, the number of true benefit-cost analyses in which net benefits is the metric of interest are few. That fact is a motivation of the Winston and Maheshrib paper.

The Winston and Maheshrib Analysis: General Issues. Clifford Winston has had a longstanding interest in transport economics, and has published widely on the topic. His body of work over the years has led him, properly so I think, to be critical of the widespread failure of governments at all levels to make decisions that raise transport efficiency and effectiveness (maximum net benefits). One can get an understanding of his frustrations by reading his opinion piece "Have Car Won't Travel; the Sober– and Sobering – Case for Privatizing Urban Transportation".[6] Having participated in transport planning myself, I certainly understand his disappointment and frustration.

In this article Winston and Maheshrib undertake an ambitious task, that of making a relatively definitive statement about the net benefits of rail transit systems in the U.S. They suggest that virtually none of these systems are justified in economic terms, arguing that the costs of such systems uniformly outweigh the benefits, with the consequent conclusion that these systems reduce the economic vitality and welfare of the nation. They argue that these systems may seem to be very beneficial to any given area, but only because many, or most, of the costs are subsidized by governments.

The question we explore here is whether they have effectively made their point. We think not, and

---

[4] Particularly biased in that regard is Thomas A. Garrett (2004), *Light Rail Transit in America: Policy Issues and Prospects for Economic Development*, Federal Reserve Bank of St. Louis (www.stlouisfed.org).

[5] Jeff Kenworthy (2008), "An International Review of The Significance of Rail in Developing More Sustainable Urban Transport Systems in Higher Income Cities," *World Transport Policy & Practice*, Vol. 14, No. 2 (www.eco-logica.co.uk); at www.eco-logica.co.uk/pdf/wtpp14.2.pdf.

[6]http://www.brookings.edu/articles/1999/04saving_winston.aspx

the recent run-up in gasoline prices is not the reason for this conclusion, although it certainly has contradicted some of their findings.

We will present arguments below that their technical analysis suffers from at least two serious defects, one statistical, and the other analytical, and as a result their conclusions cannot be sustained on the basis of their analysis. Other reviewers have reached similar conclusions.[7]

Before we examine their technical analysis, however, we note that Winston and Maheshrib generate net benefit estimates that are biased downward for reasons beyond those I present later in this paper. First, they leave out of their analysis important categories of rail transit benefits that would not show up in their statistical analysis, such as roadway and parking facility cost savings, increased traffic safety and reduced pollution emissions.[8] Second, they ignore the benefits that result from rail transit's stimulation of more compact land use which in turn increases multi-modal accessibility and generates transport cost savings through reduced per capita vehicle ownership and use which are not captured by their statistical models. Many other studies [8] [9] [10] indicate that rail transit can be a catalyst for more compact, mixed development and increased transit ridership.

Winston and Maheshrib argue that rail transit plays a declining role in the U.S. transport system, but much of their evidence reflects past trends such as housing and employment dispersion caused by massive highway investments that are now in fact reversing in major North American cities.[11] They ignore evidence of increasing demand for public transit and transit oriented development.[12] They argue that rail transit's role is declining because it only serves old central business districts, which they estimate contain only 10% of regional employment; this misrepresents the role of rail transit which often connects urban and suburban activity centers (business districts, malls, campuses and airports) and helps attract more businesses to central locations. As a result, cities with major urban rail systems tend to have a major portion of jobs, particularly higher-order jobs

---

[7] Jay Warner (2007), *Reassessing 'The Social Desirability of Urban Rail Transit Systems': Critique of Winston and Maheshrib*, VTPI (www.vtpi.org); at www.vtpi.org/warner.pdf.

[8] See Todd Litman, " Evaluating Rail Transit Benefits: A Comment," *Transport Policy*, Vol. 14, No. 1, January 2007, pp. 94-97.

[9] Jeffery J. Smith and Thomas A. Gihring, *Financing Transit Systems Through Value Capture: An Annotated Bibliography*, Geonomy Society (www.progress.org/geonomy), 2007; at www.vtpi.org/smith.pdf; originally published as "Financing Transit Systems Through Value Capture: An Annotated Bibliography," *American Journal of Economics and Sociology*, Volume 65, Issue 3, July 2006, p. 751

[10] John Schumann (2005), "Assessing Transit Changes in Columbus, Ohio, and Sacramento, California: Progress and Survival in Two State Capitals, 1995-2002," *Transportation Research Record 1930, Transit: Intermodal Transfer Facilities, Rail, Commuter Rail, Light Rail, and Major Activity Center Circulation Systems,* TRB pp. 62-67.

[11] William H. Frey (2008), *Older Cities Hold On to More People*, Census Shows, Metropolitan Policy Program, The Brookings Institution (www.brookings.edu); at www.brookings.edu/papers/2008/0710_census_frey.aspx.

[12] Reconnecting America (2004), *Hidden In Plain Sight: Capturing The Demand For Housing Near Transit*, Center for Transit-Oriented Development; (www.reconnectingamerica.org/public/tod).

that involve longer commutes, located near rail transit stations. [13]

Winston and Maheshrib argue that rail transit is inefficiently subsidized, but ignore automobile transportation subsidies such as those roadway costs not covered through user fees, subsidized parking, unpriced congestion costs, accident costs and the negative external costs of automotive emissions. Although they analyze how various transit pricing and operational changes could affect rail service cost efficiency, they fail to test the effects of efficiency-justified market reforms, such as congestion pricing, parking pricing, parking cash-out, and pay-as-you-drive vehicle insurance that would change the relative prices of transit vs. highway travel that would increase transit ridership and therefore rail transit benefits. In other words, existing market distortions that favor automobile over transit travel (subsidized parking, unpriced road space, vehicle insurance only loosely related to vehicle use) reduce transit demand below what is optimal, thereby reducing measured benefits and transit efficiencies, leading to biased statistical analysis.

The Formal Benefit-Cost Analysis.

Let us note at the outset that the basic analytical framework, that of net benefits, the difference between total benefits and total costs (NB = TB - TC) is what the authors utilize, and that is the correct general approach to the problem of evaluating the benefits and costs of a project, both prospectively and retrospectively. The simplest way to translate this into common language is that net benefit is "social profit", differing from business profit by including all consumer benefit and all social costs, such as pollution, and not just those benefits and costs which might be transacted as fares or money changing hands. Before we discuss the statistical and analytical problems with the study, we provide some discussion of the basic technique that is relevant to the following criticisms.

The Findings. We reproduce above their principal findings, their Table 5. They find that all rail systems examined except for San Francisco and Chicago are net losers, that is, total costs exceed total benefits, or TC > TB. "Consumer surplus" is their benefit measure, which is the economist's measure of consumer benefit over and above that captured by the fares riders pay. "Transit agency deficit" is the total transit agency costs minus fare collections, and "net social benefit" is the consumer surplus plus the

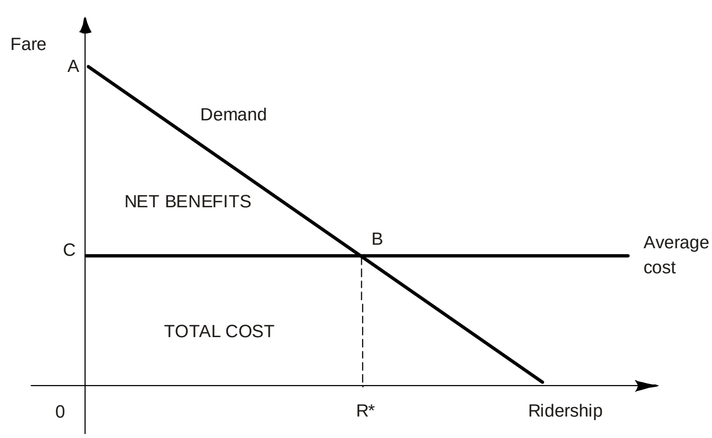

Figure 1. Net Benefits

[13] Todd Litman (2004), *Rail Transit In America: Comprehensive Evaluation of Benefits*, VTPI (www.vtpi.org); at www.vtpi.org/railben.pdf.

congestion savings minus the agency deficit, or NSB = Consumer Surplus + Congestion Savings - (Total Agency Costs - Fare Collections). We explain the basic rationale they follow with the aid of some simple diagrams.

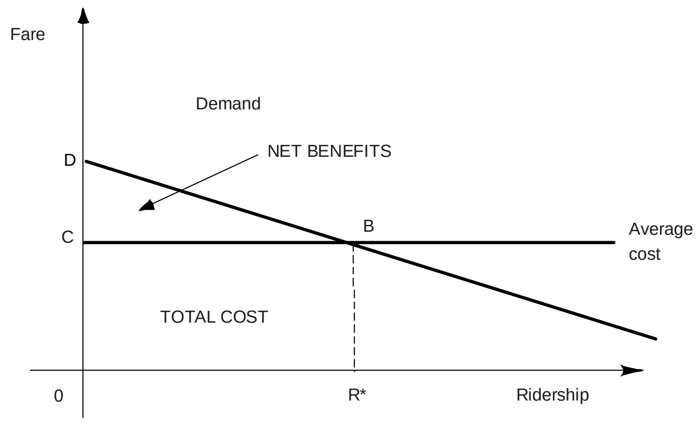

Figure 2. Lower Net Benefits

Figure 1 above captures the basic framework that the authors employ. One estimates the demand function for transit with a statistical technique called regression analysis. Note that the demand lies in price (fare) and quantity (ridership) space, and so the area under the demand curve measures the value of trips taken – this is the economist's principal measure of benefit to the riders, as the demand curve measures willingness to pay. If the fare were at the level of C (i.e., at average cost), then the fare would cover average cost, farebox collection revenue would cover costs, and ridership would be R*. We simplify the cost relationships here, just to note that the area under the average cost function measures total cost, and so net benefits is just the difference in the two areas, or area ABC. This is what the authors seek to measure, and their estimates appear in their Table 3 reproduced below. In Figure 1 consumer surplus and net benefit is also area ABC.

For our discussion here, it is important to note that two properties of the demand and cost functions can critically affect the net benefit-estimates: the height of the functions, and their slopes. We will illustrate this point with the demand function. In Figure 1 Point A on the demand curve is the vertical intercept. Should the vertical intercept be lower than that shown, say at Point D in Figure 2, then of course the net benefits will be smaller, as area DBC is smaller than area ABC in Figure 1. And should the combination of demand and cost be that shown in Figure 3, where the average cost lies above the demand curve, then the net benefits are negative, area CBFE. Figure 3 shows the case in which no level of ridership produces positive net benefits.

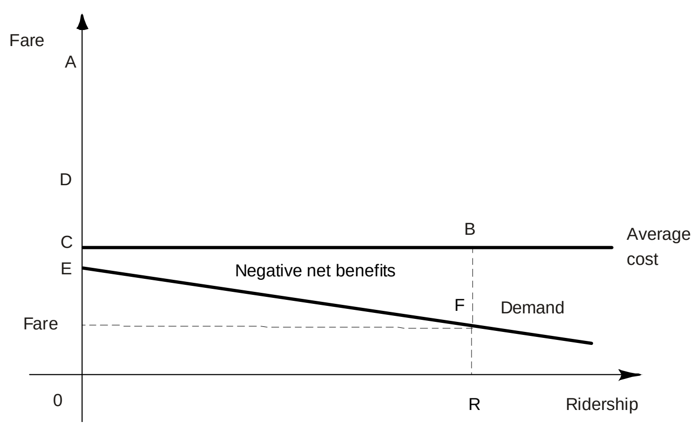

Figure 3. Negative Net Benefits

A more nuanced view of the negative net benefit situation is shown in Figure 4, where we see that the benefit-cost analysis is more complicated. Because virtually all transport systems, highways included, do not price access at cost, all modes are subsidized to one degree or another, leading some to argue, as do the authors, that there is excessive investment in the mode. There may exist some system size with a lower level of ridership for which net benefits are positive, but because of the subsidy, the higher ridership leads to excessive investment, running up the costs more than the benefits. Net benefit in Figure 4 then is area DBC minus area BGH. BGH is the excess of cost over benefit (not revenue) for the ridership between R* and R.[14] In that figure, net benefit is shown to be negative. So what economic analysis attempts to ascertain are the relative sizes of the two areas. The authors in effect argue that BGH is larger than DBC for all cities analyzed except Chicago and San Francisco, and so the net benefits are said to be negative. As we show below, they make several errors that mislead them to conclude that the situation is as in Figures 3 or 4.

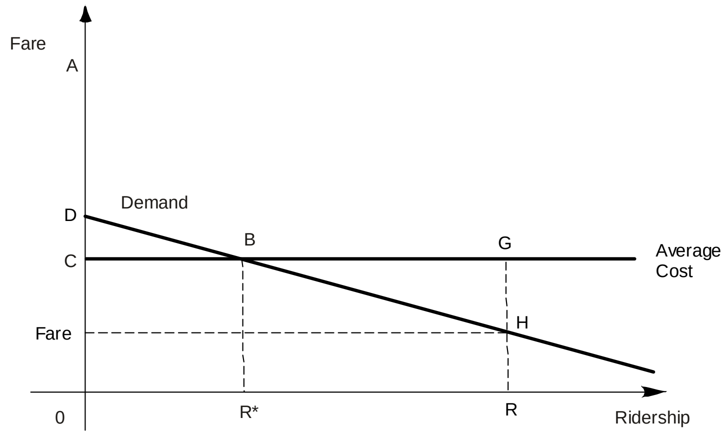

Figure 4. **Negative Nets Benefits**

The way one translates the statistical findings to this net benefits framework is fairly simple. The authors estimate a linear relationship between ridership and its causal variables. That is, they estimate an equation of the following form:

$$Ridership = a + b(fare) + c(gas\ price) + c(population) + d(income) + etc.$$

The statistical analysis determines the values of the coefficients *a, b, c, d,* and also the signs of the estimated coefficients or parameters, positive or negative. Since the net benefits analysis is expressed in fare and ridership space (price and quantity), this equation is solved (inverted) to find fare as a function of ridership, or

$$Fare = \alpha + \beta(ridership)$$

This is a straight line shown as the demand curve in the previous figures. The value of α is found by taking the average or mean values of the variables such as gas price, income, population, etc., and multiplying them by the coefficients found for the ridership equation above and then adding them together to generate α, which corresponds to the letters A, D or E in the above diagrams.

It is very important and central to these comments that if either the signs or the numerical values of the coefficients or both, are incorrect, then the net benefit calculation will be incorrect also. And the net benefit calculation may be very sensitive to alternative values of the estimated coefficients, especially if the data are very aggregated (one number per area per year) as they are

[14] It is important not to confuse revenue with benefits. Revenue is about finance, and may or may not reflect the benefits accurately. In this case, it does not, and for that reason it is important to find the demand curve and measure the area under it. One can see that fare revenue in the general case underestimates the benefits because it does not include the benefit triangle above the cost curve.

in the Winston and Maheshrib paper. So one has to be very sure about the quality of the data, i.e., that they are proper measures of the influence of interest, and resulting estimates. One also needs to do some sensitivity testing or analysis to check on how sensitive the net benefit calculation might be to a range of values. There are ways to do this, but it should be noted that the authors mention none of these, and so we conclude that they do not perform them. This in itself raises questions about the carefulness of the study.

In sum, before one can accept the author's findings, one needs to be assured that the estimates of the demand and cost functions are valid, and not biased in one direction or another. A careful econometric analysis always discusses these issues, and the normal peer review is certain to demand this of manuscripts submitted to journals for possible publication. But curiously in this article there is none of this discussion. This raises serious questions, and in fact we argue, there are reasons to think that the authors' findings are not valid and that they have underestimated the net benefits, invalidating their conclusions. This article does not show the carefulness of analysis that one would expect to support such strong conclusions and one is left to wonder about the quality of the peer review in this case.

Statistical Problems

The statistical technique employed in the study is the standard approach to econometric estimation, that of regression analysis. For those not trained in statistics, we note that in all spheres of science the way we pin down causality, that is, how the world works, is to observe a change in some independent influence or variable, and measure the change in the response variable. In controlled experiments, these changes are induced by the researcher. In natural experiments or observational situations, such as most of econometric analysis, one has to assemble the data on cause and effect variables that theorizing tells us matter and are causal, and then analyze them statistically. Multiple regression is just a formal way of doing that when there are multiple influences that affect the response variable. That variable in this study is just ridership on the demand side, and cost on the supply side.

Recalling that the authors did not have millions of dollars to conduct their study, they used published data from various governmental sources. This is standard practice and all non-experimental measurements rely on historical observations. Critical to the ability to deduce warranted implications from the statistical analysis that can be used for policy recommendations is that the data have to be, as statisticians say, "well behaved", meaning that a number of statistical properties need to be satisfied. If they are not, then the findings may be meaningless. There is a strong possibility in the Winston-Maheshrib study that the findings are meaningless, since there is one very important problem with their data that likely vitiates the conclusions they draw. We explain.

Without getting technical, since all real world variables and measurements are subject to random influences, statisticians have determined that the random errors in regression analysis on balance have to cancel out in order for the estimations to satisfy several generally accepted properties. When they don't, we say that there are systematic influences in the error term or error structure, and that causes measures of the average or mean effects, plus the measures of their variability, to be subject to biases, making the estimates either too high or too low. Conclusions based on biased measures may well be very deficient if not erroneous, and lead one to unwarranted conclusions. The first problem we mention is spurious regression.

Spurious Regression. Statisticians use a number of methods and measures to determine whether or not they have been able to pin down the cause and effect relationship in a problem. The basic two statistical measures that are always presented are a) tests of significance ("t-values") and b) the amount of variation in the dependent variable (ridership here) that is explained by the hypothesized causal variables – the R^2 or "coefficient of determination". A high R^2 suggests that one has nailed down the relationship pretty well. In the case of the Winston-Maheshrib study, this value is 0.991, nearly 100%, and the authors make note of this, but only to suggest to the reader that their findings are significant, and ignore that this high R^2 may be signaling that there are serious problems with the estimation.

In contrast to standard and accepted practice, the authors say nothing about the standard tests of significance here, and that may well be because they know that there is a serious statistical problem with the data[15]. That problem is a non-constant error mean (random errors do not cancel out) in their time series (eight years from 1993 to 2000), or what time series specialists call "non-stationarity" for short, or that there is a systematic trend in the supposedly random errors. The problem is that one can obtain statistically "meaningful" results from nonsense relationships, such as, for example, a causal relationship between total crime rates in the US between 1971 and 1991 and life expectancy of people in South Africa. There are many such examples in the literature. This problem is called "spurious regression"[16], and it does not matter, as in the Winston-Maheshrib study, that the hypothesized relationship or causal model is very realistic and well grounded in economic theory and previous findings, such as ridership being a function of transit fares and gasoline prices.

So, are the authors' data subject to this non-stationarity problem? Today, journals increasingly require authors to submit their data along with the manuscript to allow checking of the data and re-estimation of the results, a policy motivated by some well publicized cases of erroneous findings. This journal did not have such a policy. I asked Clifford Winston to share the data to allow me to check on this, but he declined, citing a Brookings policy not to share data they had paid for. I learned independently from a researcher in a position to know that Brookings in fact does not have such a policy. So the author's refusal to share the data does invite questions.

The way to determine non-stationarity is to see if there are trends, increasing or decreasing, in the data. And in fact, at least two variables in their study are characterized by trends: population and household income, and quite possibly also a third, the variable "average time to work" in some areas if congestion shows a trend, which would be the case in areas characterized by increased

[15]There are three tests that they should be reporting on: t-tests, F tests and the Durbin-Watson statistic.
[16]James D. Hamilton, **Times Series Analysis**, Princeton University Press, 1994, pp. 557-62

sprawl, and increased population.

Now this problem of non-stationarity is well known, as is the remedy: simply difference the series, that is, subtract adjacent values from one another and use the resulting differences to estimate the regression. The authors did not do this, nor do they indicate why not, and this then is another issue that raises concerns about the validity of the findings. The authors' main findings on ridership appear in the following table taken from the published article. Even before we get to discussing the implications for the estimated benefits, we note the high R^2 is inflated for two reasons: the non-stationarity problem, and the fact that the data are very highly aggregated (one observation per metro area per year). The test of the statistical significance of the entire relationship as a whole, that, all the variables in the regression equation taken together, is the F test. It is closely related to the R^2 and had they presented it, it too would be quite high, but also quite meaningless, as the non-constant error mean results in artificially inflated F tests. Again, presentation of this test value is standard operating procedure, and the fact that it is not reported has to raise concerns about the findings.

Table 3
Demand coefficients, 1993–2000* (dependent variable is billions of passenger miles)

Variable	NYC Transit	Large systems[a]	Medium systems[b]	Small systems[c]
Fare ($)	−37.1	−1.48	−3.44	−0.495
	(11.6)	(0.521)	(1.17)	(0.223)
Track density (D)**	0.544	0.544	0.061	−0.184
	(0.133)	(0.133)	(0.126)	(0.207)
Average distance between stations (η)**	−0.112	−0.112	−0.137	−0.115
	(0.057)	(0.057)	(0.047)	(0.054)
Network connectivity (Γ)**	4.94	4.94	0.013	0.206
	(1.93)	(1.93)	(0.094)	(0.163)
Network linearity (π) (greater values indicate less linearity)**	0.090	0.090	0.017	0.016
	(0.027)	(0.027)	(0.017)	(0.016)
Constant	10.6	1.43	3.56	2.92
	(3.40)	(1.95)	(1.62)	(1.52)
		Common parameters		
Rail peak-to-base ratio		(0.053)		
		(0.030)		
Bus peak-to-base-ratio		0.180		
		(0.068)		
Average number of days below 32 degrees		0.002		
		(0.001)		
Metropolitan area population (millions)		0.023		
		(0.014)		
Median household income in metro area ($thousands)		0.003		
		(0.004)		
Average regional gasoline price ($)		−0.371		
		(0.308)		
Average travel time to work (minutes)		−0.218		
		(0.105)		
Average travel time to work squared		0.004		
		(0.002)		
Year dummies		Included		
Number of observations	194	R^2		0.991

* Newey–West robust AR(1) standard errors shown in parentheses.
** We could not reject the hypothesis that the network coefficients were equal for NYC and large systems.
[a] Large systems include Washington, DC Metro, BART (SF Bay Area), Chicago Transit, MBTA (Boston), MARTA (Atlanta), SEPTA (Philadelphia) and PATH (Northern New Jersey).
[b] Medium systems include GCRT (Cleveland), MTA (Baltimore), PATCO (Southern New Jersey), San Francisco Municipal Railway, San Diego Trolley Inc., Los Angeles Metro and Metro-Dade Transit (Miami).
[c] Small systems include TriMet (Portland, OR), Santa Clara Co. Transit (San Jose), Staten Island Rapid Transit, Sacramento Regional Transit, Niagara Frontier Metro (Buffalo), Bi-State Development Agency (St. Louis), DART (Dallas), PA Alleghany Co. (Pittsburgh), RTD (Denver) and NJTransit (Newark).

Hypothesized Effects: Measurement and Data Problems

The standard way these analyses are conducted, by economists at least, is to use economic theory to specify the relationship (what causes what), and also to indicate in advance what the expected direction of the causal influence is, that is, whether the effect is expected to be positive or negative. The magnitude of the effect is found from the statistical analysis. Often the underlying theory is fairly clear about the direction of the effect, such as, prices are always inversely related to quantity demanded, or in this case ridership and fares are negatively related.

But frequently the *a priori* effect direction is not clear, and so the author will indicate that it can go either way, and offer some reasons why it might go in either direction. The problem here is that indicating no clear *a priori* direction can just be a cover hiding problems with the model or the data or both. Often if not usually, the researcher (especially economists) will have a theoretical reason to expect a positive or a negative influence, but once he or she sees that the results of the analysis do not support that, then one says that the result is “unexpected” or “surprising”, language that appears frequently in empirical papers. Or, the author may go back to the section on expected *a priori* effects and revise it with some language indicating why it might go each way.

This kind of back and forth between theorizing and findings is common in all sciences, but it is important to think through the reasons for the direction of the causality. In this case, it is clear that the authors have not spent much time on this and calls into question their findings. We explain this next, focusing on two important causal influences: gas prices and average commute time, and indicate what the implications are.

Regional gasoline prices. Generally speaking, as a microeconomic analysis (individual decision making), road trips and transit trips are substitute commuting modes, as we saw everywhere recently as gasoline prices rose. Of course, park and ride facilities combine road and transit trips, but the focus is on the main commuting mode. As substitutes, we would expect to see transit demand rise when gasoline prices rise, so the sign of the gasoline price coefficient in the regression should be positive: higher gas prices cause increased transit demand. Winston and Maheshrib do not report their data, but an examination of a gasoline price series from the Department of Energy shows that from 1993 to 2000, gasoline prices were essentially flat, falling a bit at the beginning of the period. The then current dollar prices (unadjusted for inflation) during the period lay between $1.09 and $1.47.[17] Gasoline was cheap.

But, as one of those statistical “surprises”, they find the coefficient to be negative, giving an cross-price elasticity of demand for the sample period of - 0.703 (my calculation from their reported data). This is inelastic, meaning that a 10% *increase* in gas prices will lead to a less than proportionate 7% *decrease* in transit ridership. While the magnitude of the effect seems reasonable, the sign really is unexpected, and so this qualifies as a “puzzling” and “surprising” result, although the authors do not say that.

In these circumstances, the careful researcher goes back and asks what factor or factors might explain this result, as well as scrutinizing the data to see if there is something peculiar about the data and how they are measured that might lead to the result . Often one can make a plausible conjecture. In this case, the authors conjecture that high gas prices can depress regional economic activity, and that could account for the negative coefficient, when a positive one would normally be expected. And it is true that high energy prices can have major macroeconomic effects, as we saw in the mid 1970's and early 1980's, and as we see now.

---

[17]http://www.eia.doe.gov/emeu/aer/txt/ptb0524.html

The only problem with their alternative hypothesis is that their period of 1993 to 2000 starts after the recession in 1991, and the Economic Report of the President[18] shows a prolonged economic expansion and rising employment, which the readers will recall was the period of stock market exuberance driven by the technology sector. So that explanation would not seem to work, but the authors evidently and surprisingly did not examine it further. So one is left with an unexplained finding. This well could be a result of the non-stationarity of the random errors, as discussed above, or a problem with the data. Their surmise could have easily been tested by including a measure of labor market conditions in the regressions, but they do not do this.

Now, as we discussed above, this issue is very important to the basic thrust of the paper, because in their calculation of the benefits a negative coefficient on this variable means that the estimated demand curve is shifted down (the cases illustrated in Figures 2, 3 and 4), incorrectly we argue here. That leads to a lower computed (and underestimated) value of the benefits of the transit, and certainly has to be a contributing factor to the authors' generally negative findings concerning the benefits of transit.

The data set is a panel, which means that it contains both longitudinal values (over time), and cross-sectional values (across space). There is significant variation in gasoline prices across the nation, mostly due to differences in state excise taxes on gasoline, and in the cost of fuel transport if the region does not have nearby refineries. Could this variation cause the negative relationship between transit ridership and gas prices? We do not know, as the authors neither mention nor examine it. I am unaware of any study that connects state gasoline taxes and transit ridership, but the problem with even this argument is that the generalized anti-tax political climate of the last few decades has led most states to freeze their gasoline taxes, and so the real (inflation adjusted) gasoline tax has been falling relative to the growth in income. If it were true that the states with higher gasoline taxes were also the ones with lower transit demand, that could conceivably provide the cross-sectional variation to produce the negative coefficient. That was not examined, and one could argue the converse, that those states with high taxes are the ones with large highway construction budgets, and those investments would be concentrated in congested cities, but this is just speculation. The authors could and should have examined this issue, but did not. So we are left with an implausible result, and no good explanation. One suspects a faulty model and data.

Commute time. That leads us to another important variable in the paper, average commute time. High commute times in metropolitan areas are associated with congestion. All else equal, given that road and transit trips are substitute modes for those living in the transit corridors, the predicted commuter behavioral response in the presence of frequent and expected highway congestion would be increased demand for transit, as transit lines never get congested because the headways are controlled (although the transit cars can be crowded) – thus the run times are fairly constant and quite predictable. As a result, the predicted effect of road congestion on transit demand should be positive. But again, as with gas prices, the authors instead suggest that the effect could go either way: positive or negative. They speculate that high road congestion can reduce

[18] http://www.gpoaccess.gov/eop/tables08.html

accessibility to transit stations, but offer no evidence for this hypothesis. But even if highway congestion did reduce access to park and ride facilities, it does not follow that remaining on the congested highway is a time saving strategy for journeys to work.

The authors in fact find the opposite, that increased congestion leads to lower, not higher, transit ridership. To illustrate the finding, we compute the implied cross elasticity of transit demand with respect to average commute time using their reported mean values. Using both the linear and squared terms, this elasticity is - 0.09. That is, their results indicate that a 10% *increase* in mean commute time or congestion leads to a about a 1% *decrease* in transit demand. To say the least, this is counterintuitive, and most likely is evidence of either serious data problems or faulty model specification, or both. Their explanation does appear to be an after-the-fact explanation for another "puzzling" result. And since the widely used transit demand planning models, which are based on detailed mode choice behavior, generate the benefits of transit based precisely due to reduced travel time on rail, this finding flies in the face of the established and widely used trip generation models.

A proper test of the relationship between road congestion and transit demand would require data on commuter travel in the corridors that contain *both* highways and transit lines. Congestion on roads in other corridors would not be relevant. But the data the authors employ measure the average commute time for the entire metropolitan area (one observation per metropolitan area per year), which also includes highway corridors without competing transit lines. So if congestion is severe and growing in those corridors without rail transit alternatives, as it has been, and population is decentralizing to areas without rail transit alternatives because of suburban highway investment, which it has as the authors point out [19], with the measure the authors employ, it is indeed possible to find a negative relationship between transit ridership and average commute time, that is, a correlation, but without any reason to suggest causality. This is because much or most of the congestion will be measured in non-rail transit corridors, and with the decentralization of population over time in the older transit systems to areas not served by rail, ridership will be lower. The problem here is that the commute time variable they employ is inappropriate for the analysis and leads to incorrect findings.

This has the same serious implication for the benefits calculation as the gasoline price case. A negative net sign on this pair of coefficients (for the linear and squared values) means that the demand function for transit is shifted down, and the estimated benefits reduced. But if the sign is not plausible, as we argue here, then the authors' benefit estimates are biased downward, leading to incorrect conclusions.

Mode Reliability. Related to commute time variable is the issue of the reliability of commuting via highway or transit. Commuters value not only short commutes, but reliable commutes.

[19] "The demand for rail has continued to shrink because transit networks are unable to keep up with changing land use and travel patterns that have decentralized residences and employment". Page 379. They do not mention that it is highway investments that are major drivers of this decentralization.

Reliability or risk is typically measured by the variability of the commute times, or what statisticians call the variance, yet the authors include no measure of variability of commute times, except across regions and over time. But their measure does not capture the correct dimension. We would expect that greater variability in trip times or lower reliability of highway commuting would lead to greater transit demand, and the failure to include this variable leads to the rather serious problem of "omitted variable bias". This is a major omission. Since a variable measuring relative mode reliability of transit would by hypothesized to have a positive coefficient, excluding it means that this represents a possible third reason why their demand and benefit measures would be biased downward.

Nonetheless, at this point, it is a relatively safe argument that the authors' benefit estimates are severely biased downward, and that would be one reason why they find negative net benefits for most systems. Their conclusions are erroneous as a result. The authors paid far too little attention to the quality of their data, and to whether the chosen variables or data measures were in fact well matched to the problem they sought to investigate. It is a puzzle how this got through the peer review process, on which we comment at the end.

Authors' Conclusions on the net benefits of rail systems

The measured net benefits of the systems the authors examine depend crucially on the statistical results, that is, the numerical values and signs for the parameters of the demand functions, which we noted above may well be totally meaningless in their study. If that were not enough, the authors' less than careful discussion of the net benefits of the systems adds to the doubts about the validity of the findings. We explain briefly.

Sensitivity Analysis. First, all measurements and estimates of real world relationships are subject to error. There are many types of measurement and estimation errors statisticians have identified, and so if one is to use statistical estimates to inform any decision, such as a public policy decision concerning transport investments, it is always appropriate to estimate the risks, and the consequences, of a wrong decision. "Risk" just means variable outcomes. The standard approach to this problem is to conduct a random scenario analysis, called Monte Carlo analysis, that shows the variability and the impacts on the benefit estimates. The authors do not do this.

Additionally, a careful benefit-cost analysis always presents a discussion of those variables that the authors have not been able to include, usually for lack of data availability. Such a discussion includes a listing of these excluded variables and the likely impacts or implications of the exclusion. This is critical, because of the omitted variable bias mentioned above, and we illustrate with the variable of highway lane miles.

This "omitted variable" problem seems to be present in this study. This potentially serious omission is that they do not include highway expansions (lane miles) in the variable set. Road traffic and transit ridership are jointly determined, in which the time and money cost of one mode affects the demand for the substitute mode. In the period they study, 1993 to 2000, there was important economic growth throughout the country, populations continued suburbanizing, and urban highway capacity was added, but only in a few new systems has track capacity been expanded. More highway lane miles leads to lower transit ridership as it can reduce road congestion in the short run (a period of several years), thus reducing the time and money costs of highway commutes. So there is a negative influence that appears in the ridership numbers, but no independent causal variable designed to capture and measure it. This means that one or more of the estimated parameters are in error, either the sign, or the magnitude, or both. In this case, the influence of this omitted variable may be weakly captured in the variable "average time to work", but we cannot know without a complete analysis of the relationships without the variable. In a multiple regression it is a very tedious exercise to work out the effect of this bias analytically, that is, by hand.

Further, as noted, road traffic and transit ridership are jointly determined. That raises still another problem, called simultaneous equation bias, which also has the effect of generating biased values for the responsiveness of ridership to variation in fares and other influences such as gasoline prices and congestion.

This matters to the benefit-cost analysis in this way: a relevant question for transport investments is what would have happened to transit ridership had the additional investments in highway expansions not taken place. Multiple regression always allows one to provide some answers to that question, which would be critical to the conclusions that the authors draw.

Option Value. Another dimension of careful benefit-cost analysis is attention to those benefit and cost dimensions that are not captured with the techniques they employ. For example, people generally like to have options when making choices, both now and in the future. So in general even those who currently choose to commute by auto would be generally willing to pay some positive amount to ensure that other transit options remain open to them should they be unable or unwilling to drive at some future time, due to such factors as car repairs, highway construction, incapacitating health conditions, *etc*.

Logically then economists call this "option value", and this value is not picked up in the demand function that the authors estimate. There are many options markets and have been the object of much study. There are suggested techniques for the transit case[20], which would not have required much additional work to apply in this study. But at a minimum the issue should have been discussed, with the observation that those benefits that they did estimate are too low because of the exclusion of option value.

Further, they could have asked the question about how large option value would have to be in order for the net benefits in their study to be positive. In this case, the estimate would be too large because, as we have argued above, their net benefit estimates are biased downward, and are too low.

[20]See the Transit Cooperative Research Program report 78, "Estimating the Benefits and Costs of Public Transit Projects: A Guidebook for Practitioners" Transportation Research Board, National Academy of Sciences, 2002.

Conclusions

In summary, while at first blush Winston and Maheshri's study seems carefully done, in fact it contains significant biases and omissions. It only considers a minor portion of rail transit benefits and ignores the indirect social benefits that result when rail transit provides a catalyst for more compact development. It overlooks trends that are increasing demand for high quality public transit service and transit oriented development. As a result, rail transit investments tend to provide far greater total benefits than the study assumes.

There are some relatively elementary statistical issues that were not explored, or at least, no reasons were presented indicating why they were not, and so on this basis alone it is inappropriate and unwarranted for the authors to reach the conclusions they do: that virtually all of the rail systems are net losers.

In fact, this seems less than a carefully done analysis. When asked why he did not present the statistical tests mentioned above, Winston's reply was that the data series is too short to present these problems. That is possible, but not one that the profession would take on faith, and so evidence should be provided. The tests should have been done and the fact that they were not properly and normally raises questions about the validity of the findings as well as about potential biases.

Finally, how did this get through the review process? That is for the journal editor to answer, but having participated in conference volumes before, the answer quite likely is that the normal review process was not applied here, but something less exacting. At least this manuscript would not have made it by this reviewer without a very substantial "revise and resubmit" recommendation to the editor. We have to conclude that this study has multiple serious shortcomings and should not be used to draw conclusions about the direction of investments in the nation's transport system, especially not in this time of rising energy prices.