

Letter to the Editor, *JAPA*.

Credibility Problem for Transit: Response to Millar

The *JAPA* Summer 2005 issue contains a letter from the President of the American Public Transportation Association, William Millar, regarding our article "How (In)accurate Are Demand Forecasts in Public Works Projects? The Case of Transportation," printed in the Spring issue. We concluded that transit ridership forecasts are frequently and highly overestimated, including those for projects in the U.S. Millar challenges this conclusion and calls for readers to disregard our "specious article." We welcome this opportunity to clarify the points raised by Millar.

We will respond to Millar's comment point by point.

Millar says that the U.S. transit ridership forecasts in our article are "very old" and that notable forecasting improvements have been made in the last two decades that are not included in our article. Millar is wrong on both counts. Our article includes the latest U.S. transit ridership forecasts for which data on outcomes are available (19 forecasts made in the 1980s and 1990s). Information on these forecasts was made available by the Federal Transit Administration (described on p. 136 of the article). Our article also contains older forecasts in order to test whether forecasting accuracy has improved over time. Such tests, documented in our article, show that two thirds of new transit ridership forecasts are still off by more than 20%, and many are off by much more. This is a change from being very, very wrong to being very wrong. To call this a "notable improvement," as does Millar, seems to us quite a stretch. The FTA's conclusion that "Risk of large errors still remains" is more in line with the evidence.

Millar is also mistaken when he says that only 27 projects in our study were public transportation, and that this is not an informative sample. Our article in fact covers 46 public transportation projects, of which 27 are U.S. projects. This is the first time such a sample has been established that is large enough to allow statistically significant conclusions. Our conclusion that transit ridership forecasts are frequently and highly overestimated is particularly well founded for the U.S., because the number of observations is largest for this country.

Millar quotes examples from St. Louis, Salt Lake City, and Portland of ridership forecasts that were exceeded by actual ridership as evidence against our conclusion that transit ridership forecasts are frequently and highly overestimated. Like Millar, we're gratified that such examples exist, but they are obviously rarities given that 8 of 10 projects have ridership shortfalls. Moreover, these examples are actually included in the FTA data covered in our article. Thus, of the 27 U.S. transit projects in our sample, Millar selects a few atypical ones and presents them as evidence against the general argument. For reasons of decorum, we will pass over the word statisticians use to describe such handling of data.

Millar criticizes us for not taking into account positive development effects of transit, for instance as reflected in increased real estate values. We agree that such effects exist, but the critique is misguided, since the aim of our study was to measure inaccuracy in ridership forecasts and not development effects. However, if transit has development benefits, these must be expected to be roughly proportional to ridership, so that if ridership has been overestimated, so must have the development benefits, which means that the case for proceeding with the project was exaggerated.

Finally, Millar is critical of our research for using the first year of operations to determine whether results met forecasts. He argues that it takes considerable time for

development and travel patterns to change and, therefore, the first year of operations is a poor measure of accuracy (this is the so-called "ramp up" argument). A study of this assertion, which was published in the July 2005 issue of *Transportation Research A* ("Measuring Inaccuracy in Travel Demand Forecasting: Methodological Considerations Regarding Ramp Up and Sampling," by Bent Flyvbjerg), shows Millar to be wrong.

The short answer to the ramp up argument is that if it is significant, it is presumably already reflected in ridership modeling, making forecasts more accurate both for the first year of operations and later years. It turns out, however, that ramp up is typically not included in forecasting models, with good reason. Ramp up appears not to be significant for most projects. Among projects for which available data on actual and forecasted traffic cover more than one year after operations began, projects with lower-than-forecasted traffic during the first year of operations also tend to have lower-than-forecasted traffic in later years. Thus, using first-year ridership to determine forecast accuracy appears unlikely to mistakenly identify as underperforming projects that would appear successful in a later year. Actual traffic apparently does not quickly catch up with forecasted traffic for this type of project, and sometimes it never does, so that assuming ramp up for such projects would lead to less accuracy in ridership forecasts, not more.

We understand it is William Millar's job to promote the interests of public transportation and we agree that transit makes a critical and necessary contribution to the quality of urban life. We want to suggest, however, that Millar does public transportation a disservice when he misrepresents and ignores research findings he doesn't favor. Worse still, Millar encourages readers to disregard a study that contains data and a method that can effectively improve accuracy in transit ridership forecasts and thus improve transit planning. If transit decision making were less distorted by spin, the

public would benefit by better investments, and Millar's industry would gain by having greater support from the public.

The American Planning Association advocates using the reference class forecasting method we propose in our article ("APA encourages planners to use reference class forecasting in addition to traditional methods as a way to improve accuracy," APA Press Release, April 7, 2005, available at <http://www.planning.org/newsreleases/2005/ftp040705.htm>). We suggest that U.S. planners follow this advice, like their colleagues in the U.K., and see Millar's critique for what it is, a lobbyist's attempt to cover up the unpleasant fact that for decades transit forecasting and decision making have been self-servingly misleading.

Bent Flyvbjerg, Mette K. Skamris Holm, and Søren L. Buhl

Aalborg University, Denmark