Transportation Research Part A 96 (2017) 101–118

Contents lists available at ScienceDirect

Transportation Research Part A

journal homepage: www.elsevier.com/locate/tra

Validation of aggregate reference forecasts for passenger transport

Matts Andersson [a,b,*], Karin Brundell-Freij [a], Jonas Eliasson [b]

[a] WSP Analysis & Strategy, Sweden
[b] Department for Transport Science, KTH Royal Institute of Technology, Sweden

ARTICLE INFO

Article history:
Received 5 June 2016
Received in revised form 15 December 2016
Accepted 19 December 2016
Available online 4 January 2017

JEL classificationCodes:
R41
R42

Keywords:
Demand modeling
Forecast
Transport
Accuracy
Validation

ABSTRACT

We have compared Swedish national forecasts for passenger transport produced from 1975 to 2009 with the actual outcomes, and we found substantial differences between forecasts of passenger kilometers by mode and actual outcomes. In forecasts produced since the early 1990 s, road and air traffic growth rates have generally been overpredicted. Aggregate railway growth has been fairly accurate, but commercial long-distance railway growth has been overpredicted, and the growth of subsidized intra-regional railway travel has been underpredicted (following vast unanticipated supply increases).

Focusing on car traffic forecasts, we show that a very large share of forecast errors can be explained by input variables turning out to be different than what was assumed in the forecasts. Even the original forecasts are much closer to actual outcomes than simple trendlines would have been, and once the input assumptions are corrected, the forecasts vastly outperform simple trendlines. The potential problems of using cross-sectional models for forecasting intertemporal changes thus seem to be limited. This tentative conclusion is also supported by the finding that elasticities from the cross-sectional models are consistent with those from a time-series model.

1. Introduction

Transport forecasts play an important role in policy design, decision-making, and public debate. The question addressed in this paper is how accurate forecasts for passenger transport are and to what extent forecast errors can be explained by wrong assumptions about input variables. When discussing the validation of transport forecasts, it is useful to distinguish between *reference* forecasts, which predict transport volumes in a future year in a baseline scenario, and *policy* forecasts, which predict the effects of some policy by comparing "do-nothing" and "do-something" scenarios. Such "policies" can be any intervention in the transport system, including infrastructure investments, changes in prices or taxes, new regulations, and so on. In practice, a *project* forecast – such as the forecasted traffic volume on a planned road – is a mixture of the two in which a reference forecast gives the baseline traffic volumes in the do-nothing scenario (which is usually several years into the future), and a policy forecast predicts how traffic volumes will be changed by the new road. When analyzing forecast accuracy, however, it is useful to distinguish between the two for several reasons. In this paper, we analyze reference forecasts.

* Corresponding author at: Department for Transport Science, KTH Royal Institute of Technology, Sweden.
E-mail address: matts.andersson@wspgroup.se (M. Andersson).

http://dx.doi.org/10.1016/j.tra.2016.12.008
0965-8564/

Reference forecasts tend to form our picture of the future and our intuitive view of what investments and policy measures we believe will be necessary. The importance of reference forecasts makes them debated. For example, in Norway they are often accused of underestimating road traffic volumes thus leading to underestimations of road investment benefits (Kjerkreit and Odeck, 2009). In Sweden, voices in the debate have claimed the opposite (Ljungberg, 2014).

Sweden has a long tradition of producing (semi-)official national reference forecasts that form the baseline scenario for long-term national transport planning. This paper compares eight Swedish national forecasts for passenger traffic made between 1975 and 2009 with actual outcomes. Forecasts for car, rail and domestic air are analyzed. Bus travel is not analyzed since the statistics are to crude for this to be meaningful (only based on vehicle mileage). We also try to explain to what extent differences between forecasts and outcomes can be explained by wrong assumptions about input variables such as fuel prices and economic growth. Here, the focus is primarily on car traffic because assumptions about supply data (fares, service frequencies, etc.) for rail and domestic air are not documented well enough to make this possible.

The models used to produce the forecasts have varied over time, but they have all been variants of 4-step models producing detailed forecasts on network and origin-destination levels, and they have all been estimated on cross-sectional data. The present paper focuses on how well the models have managed to predict aggregate numbers of total passenger kilometers by mode some time into the future (typically more than a decade). Using models estimated on cross-sectional data to predict future transport patterns implicitly assumes that cross-sectional relationships between variables and travel behavior are the same as the corresponding inter-temporal relationships. For example, if income levels increase over time, this is assumed to affect travel behavior in a way that corresponds with observed cross-sectional differences across income groups. Analogous assumptions are (implicitly) made for the effects of changes in car ownership, urbanization, and so on. A central, long-standing question in the transport forecasting literature is to what extent models estimated on cross-sectional data can be used to produce reference forecasts stretching decades into the future. One of the purposes of the present paper is to explore this issue. Focusing on aggregate reference forecasts (future passenger kilometers) produced by cross-sectional models might thus be seen as almost unfair as we focus precisely on what should be expected of the weakest point of such models. Still, the question is important because reference forecasts are most often produced by cross-sectional models. This practice is common due to the need for detailed forecasts for project analyses such as infrastructure investments. So far, it is only models based on cross-sectional data that are able to produce such detailed forecasts, and time series-based models are almost never able to produce network-level forecasts working instead only at an aggregate level.

The remaining paper is structured as follows. A literature review is given in Section 2, data sources and the background to the forecasts we analyze are presented in Section 3, comparisons between forecasts and outcomes are presented in Section 4, an analysis of the effect of input errors for car forecasts is presented in Section 5, and the concluding remarks are given in Section 6.

2. Literature

When discussing forecast validation, a distinction between reference forecasts and policy forecasts needs to be made since conclusions regarding strengths and weaknesses might be very different. For example, if (hypothetically) reference forecasts tend to underestimate the growth of a particular mode, it does not necessarily follow that forecasts of the effects of an investment in that mode are underestimated. In particular, the distinction is important if the purposes are to improve forecasting models and to assess to what extent, under what circumstances, and for which applications models can be trusted. Often, however, this distinction is blurred; in particular, demand forecasts for any particular project are necessarily a mixture of the two, and only careful econometric work can disentangle errors in the reference forecast from errors in the policy forecast. Consequently, most forecast validation studies do not explicitly separate the two.

Most of the forecast validation literature deals with project-specific forecasts. There are, however, a few studies that have followed up on reference forecasts on a more or less aggregate level. SIKA (2005) compared Swedish forecasts made from 1975 to actual outcomes and concluding that forecasts had difficulties in catching trend breaks. Odeck (2013) followed up road traffic forecasts for Norway (both for different counties and the country as a whole) for the period 1996–2008 and found that the accuracy of the national forecasts was improved after a major model revision in 2001. National forecasts were unbiased after the model revision, but there were still variations among regions and time periods. The Department of Transport in the UK has carried out back-casting evaluations of their national transport model by using the model to make "predictions" for the years 1976 and 1991 and comparing these to national travel surveys from those years (Gunn et al., 2006). The predictions were for the number of trip-ends and for mode/destination choice, and they did not use the entire model to make transport forecasts. The model predicted the number of trips correctly for 1991, but overestimated it by 15% for 1976. Trip-making in the London area was over-predicted by around 20% for both years (which was offset by under-predictions in the rest of the country). The authors found that the model had performed well in predicting back to 1991, but they concluded that the model had predicted little change in a situation in which there had been little change in the factors affecting mode and destination choice at an aggregate level.

Nicolaisen (2012) presented an ex-post evaluation of demand forecast accuracy for do-nothing alternatives based on an empirical study of 35 non-tolled road projects in Denmark and England. The results showed a tendency for systematic over-estimation of travel demand in the do-nothing alternatives, on average by 7%. This was in contrast to the systematic under-estimation of travel demand observed in the "do-something" forecasts. The study also found that the variation in do-nothing