# Understanding the trade offs between DTA models realism and robustness: the impact of spillback modeling.

Stephen D. Boyles[*], Natalia Ruiz Juri[†]

December 15, 2017

## 1 Introduction

Bringing advanced models to practice often involves moving from simpler, stylized models to more realistic ones by relaxing assumptions. While more realistic models are *ceteris paribus* preferable to less realistic ones, realism is rearely the only criterion for model selection. Realism must be balanced against tractability, availability of data for calibration and validation, mathematical regularity, and robustness to errors in the input data. Practitioners must make modeling choices taking all these factors into account, given their relevance in a specific application.

This paper specifically examines the trade offs between model realism and robustness to input errors, in the context of dynamic network loading for dynamic traffic assignment (DTA) implementations. We focus on the model capability to represent congestion spillbacks, which are responsible for a significant portion of teh observed congestion in real transportation networks.

While capturing splillbacks is clearly a critical feature of a practical DTA model, spillbacks may lead to excessive congestion and gridlocks, which typically result in poorer convergence and less stable models. The latter is often observed in future year models for which inputs are inaccurate, and traffic control at intersections can be only coarsely approximated. Further, in practice, input data are never known with exact precision due to measurement, estimation, and forecasting errors. Therefore, it is not obvious that the more "realistic" model is preferable to one which is more robust to these errors. Rather, one or the other model may be preferable depending on (1) the level of uncertainty in the input data, (2) the relative sensitivities of the two models, (3) and the magnitude of the error introduced by ignoring spillback.

In this paper we explore the impacts of the error introduced by ignoring queue spillback (*systematic error*) to the error introduced from incorrect input data (*random error*) when spillbacks are explicitly modeled. Our approach includes conceptual tests in a simple setting, and numerical examples in large real-world networks.

Preliminary findings in a simple freeway interchange suggest that in cases of low data uncertainty the model with spillback is clearly preferred; in the case of high data uncertainty, the no-spillback model often has the lower expected absolute error. The full paper will include similar analyses on larger-scale urban networks, and we will explore practical solutions to address issues related to unrealistic spillback impacts in future year models. The following sections illustrate the concepts previously described through a simple example.

[*]Assistant Professor; Department of Civil, Architectural and Environmental Engineernig; The University of Texas at Austin
[†]Research Associate; Center for Transportation Research; The University of Texas at Austin

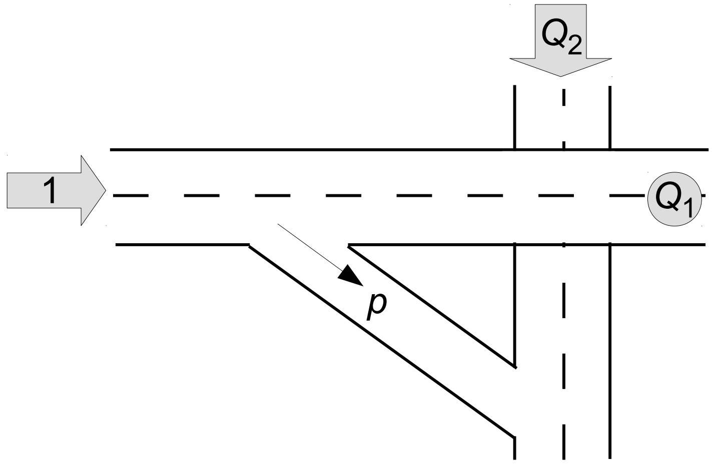


Figure 1: Freeway interchange for demonstration; mainlines have capacity 1 and ramp has capacity 1/2.

# 2 Freeway interchange scenario description

Consider the freeway interchange shown in Figure 1. The two mainlines have equal capacity, and the ramp connecting them has half the capacity of the mainlines. For simplicity, assume that units are chosen such that the capacity of each mainline is equal to one, and that the inflow rate on the horizontal mainline is also one. There are two model parameters which must be estimated: the proportion $p$ of flow on the horizontal link choosing the ramp, and the flow $Q_2$ on the vertical mainline link. In this demonstration both are assumed stationary in time. We wish to estimate the steady-state flow rate on the horizontal link downstream of the diverge, shown by the circular detector in Figure 1, after any initial transient conditions have subsided.

We use the standard merge and diverge equations in dynamic network loading, assuming that in highly congested conditions the merge allocates flow to approaches proportionate to their capacity, and that diverges respect the first-in, first-out principle. Since the inflow to the ramp is initially $p$, if $p + Q_2 \leq 1$ the merge is uncongested and all flow can move freely. If $p + Q_2 > 1$, three cases are possible: if $p < 1/3$, a queue forms on the vertical link but not the ramp; if $Q_2 < 2/3$, a queue forms on the ramp but not the vertical link; and if $p \geq 1/3$ and $Q_2 \geq 2/3$, queues form on both approaches. In a point queue model, the presence of queues at the merge has no bearing on behavior at the diverge, and the total outflow from the diverge is given by $\min\{1, 1/(2p)\}$, with the detector registering

$$Q_1^{NS} = (1-p)\min\{1, 1/(2p)\}\,. \tag{1}$$

In the case of spillback, equation (1) only holds if there is no queue on the ramp. Otherwise, at steady-state, the ramp outflow (and thus its inflow) is given by $1 - Q_2$ if the queue is only on the ramp, and by 1/3 if queues exist on both the ramp and vertical link. The corresponding flow at the detector is obtained by multiplying the ramp flow by $(1-p)/p$. These results are summarized in Figure 2, and are denoted by the mapings $Q_1^{NS}(Q_2, p)$ and $Q_1^S(Q_2, p)$ for the no-spillback and spillback cases, respectively. Full derivations are omitted here for reasons of space.

# 3 Procedure

The possible values of inputs $Q_2$ and $p$ lie within the unit square $[0, 1]^2$. Within this range, twenty evenly-spaced values of $Q_2$ and $p$ were combined to produce four hundred scenarios for analysis. For each of these

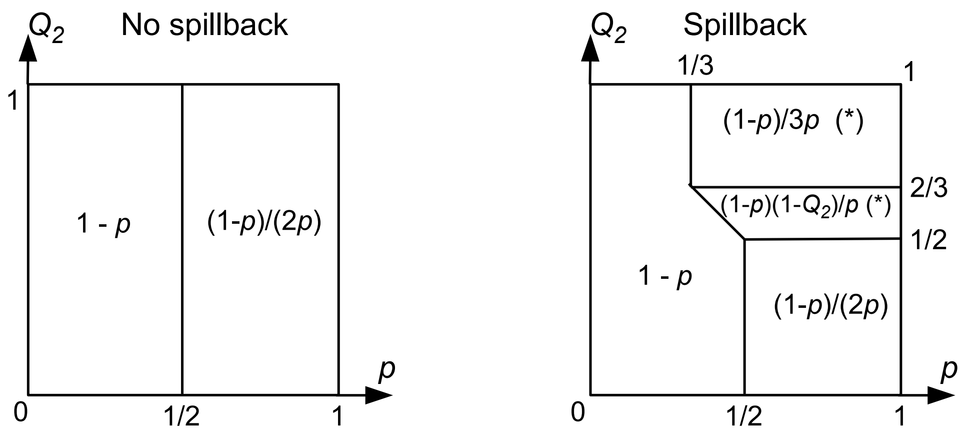


Figure 2: Steady-state $Q_1$ values for no-spillback and spillback cases. Asterisks indicate regions where a ramp queue restricts inflow.

scenarios, the following procedure was performed.

1. Let $\hat{Q}_2$ and $\hat{p}$ denote the values corresponding to this scenario. These are assumed to be the true values of these parameters, corresponding to a true flow rate of $Q_1^S(\hat{Q}_2, \hat{p})$. (That is, the spillback model is presumed completely accurate if given the true $Q_2$ and $p$ values.)
2. Generate $n$ sampled values of $Q_2$ and $p$, using independent normal distributions with respective means $\hat{Q}_2$ and $\hat{p}$, and a provided standard deviation.
3. For each sample, the error associated with the no-spillback model is calculated as

$$\epsilon^{NS} = \left|Q_1^{NS}(Q_2, p) - Q_1^S(\hat{Q}_2, \hat{p})\right|, \tag{2}$$

   while the error associated with the spillback model is

$$\epsilon^{S} = \left|Q_1^{S}(Q_2, p) - Q_1^S(\hat{Q}_2, \hat{p})\right|. \tag{3}$$

4. Based on the sampled values of $\epsilon^{NS}$ and $\epsilon^S$, calculate the additional expected error in the no-spillback model to be

$$\delta = E[\epsilon^{NS} - \epsilon^S], \tag{4}$$

   along with the standard deviation $s$ of this difference.
5. Calculate the $t$ score $t = \delta/(s/\sqrt{n})$.

If the resulting $t$ score is greater than a specified positive critical value, the error in the no-spillback model is significantly greater than that in the spillback model, and the spillback model is to be preferred. If it is smaller than a negative critical value, the error in the no-spillback model is significantly less than that of the spillback model, and the no-spillback model is preferred. Otherwise, there is no significant difference in the errors produced by the models.

# 4 Discussion and Conclusions

Figure 3 presents the results of these simulations for three cases: when the standard deviation of the sampled $Q_2$ and $p$ values was small (0.01), moderate (0.1), and large (0.25). A sample size of $n = 2500$ was used for

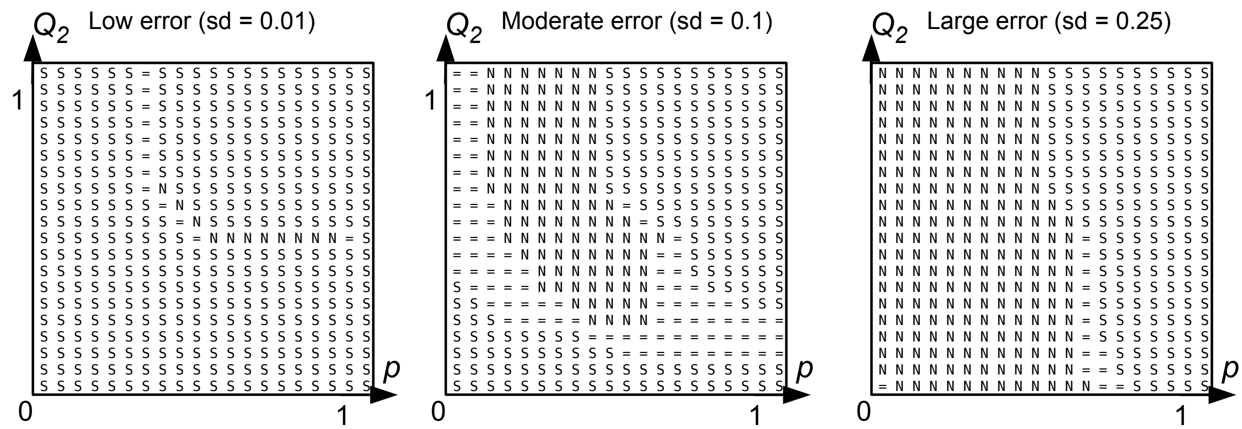


Figure 3: Model with less error for each scenario, if significant ($p = 0.05$).

each scenario, and critical values of $\pm 1.96$ were used for the statistical test, corresponding to 5% significance. In this figure, an S denotes that the model with spillback produces less expected error, NS denotes that the no-spillback model produces less expected error, and '=' denotes no statistically significant difference in errors. Owing to the sample size, the $t$ scores were typically quite large, averaging $+87.4$ across all scenarios where the spillback model was preferred, and $-13.9$ across all scenarios where the no-spillback model was preferred.

When input error is small, the model with spillback is almost always preferred, while the no-spillback model often produces less expected error when input errors are large, specifically when the diverge proportion is small. This demonstrates that the no-spillback model may be preferable to the model with spillback under certain circumstances: even though it contains systematic error, it is more robust to random error.

The full paper examines this idea in the context of a dynamic traffic assignment model of a realistic urban area. Specifically, this larger experiment will allow for equilibrium route choice, which may act as a "restoring mechanism" against errors in the input data. The analysis will explore whether, and under what conditions, including spillback can increase error.